# Nonlinear Dynamics and Chaos

Prof. George Haller*[1], Trevor Winstral[2], and Bálint Kaszás[1]

[1] *Chair in Nonlinear Dynamics, Institute for Mechanical Systems, ETH Zürich*
[2] *Department of Mathematics, ETH Zürich*

2022

151-0532-00L, 151-0530-00L, Github Repository
*Website, Email

# Contents


**I Nonlinear Dynamics and Chaos 1** **5**

**1 Introduction** **7**

**2 Fundamentals** **13**

2.1 Existence and uniqueness of solutions . . . . . 13
2.2 Geometric consequences of uniqueness . . . . . 16
2.3 Local vs global existence . . . . . 17
2.4 Dependence on initial conditions . . . . . 19
2.5 Dependence on parameters . . . . . 26

**3 Stability of fixed points** **33**

3.1 Basic definitions . . . . . 33
3.2 Stability based on linearization . . . . . 39
3.3 Review of linear dynamical systems . . . . . 40
3.4 Stability of fixed points in autonomous linear systems . . . . . 42
3.5 Stability of fixed points in nonlinear systems . . . . . 46
3.6 Data-driven linear modeling of dynamical systems . . . . . 60
3.7 Lyapunov's direct (second) method for stability . . . . . 62

**4 Bifurcations of fixed points** **73**

4.1 Local nonlinear dynamics near fixed points . . . . . 73
4.2 The center manifold . . . . . 75
4.3 Center manifolds depending on parameters . . . . . 81
4.4 Normal forms . . . . . 83
4.5 Bifurcations . . . . . 86
4.6 Codimension 1 bifurcations . . . . . 88

**5 Nonlinear dynamical systems on the plane** **99**
5.1 One degree of freedom conservative mechanical systems . . . 99
5.2 Global behavior in two dimensional autonomous dynamical systems . . . 103

**6 Time-dependent dynamical systems** **109**
6.1 Nonautonomous linear systems . . . 109
6.2 Time-periodic nonlinear systems . . . 110
6.3 Averaging . . . 112
6.4 The Harmonic Balance Method . . . 119

**II Nonlinear Dynamics and Chaos 2** **125**

**7 Indroduction to chaotic dynamics** **127**
7.1 Consequences of hyperbolicity . . . 130
7.2 Dynamics near the homoclinic tangle & Smale's horseshoe map . . . 137
7.3 Dynamics of Smale's horseshoe map on the invariant set $\Lambda$ . . . 141
7.4 Generalization of chaotic dynamics . . . 153

**8 Introduction to Hamiltonian Systems** **161**
8.1 Properties of Hamiltonian systems . . . 167
8.2 Volumes under Hamiltonian flow maps . . . 171
8.3 Consequences of area preservation . . . 174
8.4 Stability of fixed points in Hamiltonian systems . . . 182
8.5 Stability of Hamiltonian fixed points: normal forms . . . 187

**9 Normally hyperbolic invariant manifolds** **191**
9.1 A crash course on differentiable manifolds . . . 191
9.2 Theory of normally hyperbolic invariant manifolds . . . 200
9.3 Geometric singular perturbation theory . . . 207
9.4 Existence of invariant manifolds . . . 218
9.5 Spectral submanifolds and model reduction . . . 223

# Part I

# Nonlinear Dynamics and Chaos 1

# Chapter 1

# Introduction

First we shall introduce the most important characters in our following exploration. The ideas and definitions here will be recurring regularly as we examine them from different perspectives and using different tools. The content covered by this course can be found in the following books. For further details on some of the results, we recommend consulting these.

- J. Guckenheimer & P. Holmes, Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields [Guckenheimer and Holmes, 1990];
- F. Verhulst, Nonlinear Differential Equations and Dynamical Systems [Verhulst, 1989];
- V. I. Arnold, Ordinary Differential Equations [Arnold, 1992];
- S. Strogatz, Nonlinear Dynamics and Chaos [Strogatz, 2000].

**Definition 1.1** (Dynamical System (DS))**.** A triple $(P, E, \mathcal{F})$, with

- $P$ : the phase space for the dynamical variable $x \in P$,
- $E$ : base space of the evolutionary variable (e.g. time) $t \in E$,
- $\mathcal{F}$ : the evolution rule (deterministic) which defines the transition from one state to the next.

The two main types of evolutionary variable spaces are

(i) Discrete dynamical systems (DDS) $t \in E = \mathbb{Z}$ with trajectory $\{x_0, x_1, \ldots\}$,

(ii) Continuous dynamical systems (CDS) $t \in E = \mathbb{R}$ with trajectory $\{x_t\}_{t\in\mathbb{R}}$.

Corresponding to these there are various types of evolution rules

(i) In a DDS we have iterated mappings

$$\boxed{x_{n+1} = F(x_n, n).} \tag{1.1}$$

If there is no explicit dependence on $n$, i.e. $\frac{\partial F}{\partial n} = 0$, then

$$\boxed{x_{n+1} = F(x_n) = F(F(x_{n-1})) = \underbrace{F \circ \ldots \circ F}_{n+1 \text{ times}}(x_0) = F^{n+1}(x_0).} \tag{1.2}$$

*Example* 1.1 (Cobweb diagram of a one-dimensional DDS). In such cases and in one-dimensional problems, a simple way to analyze the behavior of the system is the so-called *cobweb* diagram. We may plot $x_{n+1}$ as a function of $x_n$, as demonstrated in Fig. 1.1. The image of an initial condition $x_0$ lies on the graph at $x_{n+1} = F(x_0)$. We can also compute the next iterate by horizontally projecting the point $(x_0, F(x_0))$ to the diagonal line defined by $x_{n+1} = x_n$. Following the projection of this point to the horizontal axis $(x_n)$ we find the intersection with the graph at the point $(x_1, F(x_1))$. It follows that fixed points on the cobweb diagram correspond to the intersection of the graph of $F$ with the diagonal line $x_{n+1} = x_n$.

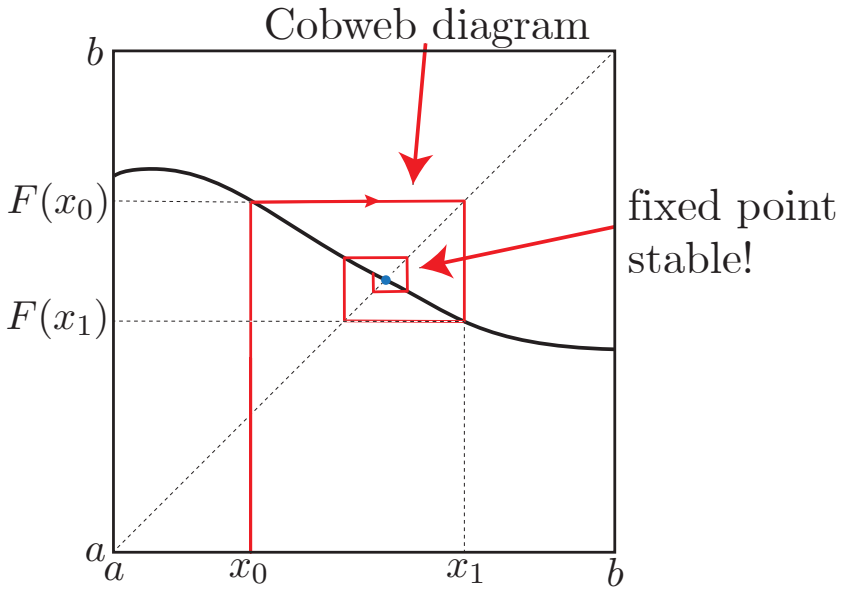


Figure 1.1: Analysis of a one-dimensional system defined on the interval $x \in [a, b]$ using the cobweb diagram

(ii) In a CDS we have a first order system of ordinary differential equations (ODE)

$$\boxed{\dot{x} = f(x, t)} \tag{1.3}$$

for $x \in P$ and $t \in E$. This yields the initial value problem (IVP):

$$\begin{cases} \dot{x} = f(x,t) \\ x(t_0) = x_0 \end{cases} \tag{1.4}$$

Assuming there exists a unique solution $\varphi(t; t_0, x_0)$ with $\dot{\varphi} = f(\varphi, t)$ and $\varphi(t_0) = x_0$, then the following flow map is well defined

$$\boxed{F_{t_0}^t(x_0) := \varphi(t; t_0, x_0).} \tag{1.5}$$

Geometrically, this solution can be viewed as a trajectory in phase space (cf. Fig. 1.2).

Figure 1.2: Trajectory of a continuous dynamical system. The RHS is given by f(x,t), which is the tangent vector to this curve at the point $x$ at time $t$.

Such an $F_{t_0}^t$ has the properties

(a) $F_{t_0}^t$ is as smooth as $f(x,t)$,

(b) $F_{t_0}^{t_0} = I$ and $F_{t_0}^{t_2} = F_{t_1}^{t_2} \circ F_{t_0}^{t_1}$,

(c) $\left(F_{t_0}^t\right)^{-1} = F_t^{t_0}$ exists and is smooth.

Properties (b) and (c) together are called the group property. A special case of continuous dynamical systems is the autonomous system.

$$\boxed{\dot{x} = f(x).} \tag{1.6}$$

The autonomy of a system implies

$$x(s, t_0, x_0) = x(\underbrace{s - t_0}_{t}, 0, x_0) \overset{!}{=} x(t, x_0). \tag{1.7}$$

The induced flow map in this case is the one-parameter family of maps

$$\boxed{F^t = F_0^t : x_0 \mapsto x(t, x_0).} \tag{1.8}$$

*Example* 1.2 (Logisitic Equation). For a resource-limited population, we have the following dynamical system for $a > 0$, $b > 0$, and the population $x \in \mathbb{R}_+ \cup \{0\}$

$$\dot{x} = ax(b - x). \tag{1.9}$$

In this case we have $E = \mathbb{R}$ and $\mathcal{F} = \{F^t\}_{t=-\infty}^{+\infty}$. This system has globally existing unique solutions (see later). We may analyze the behavior of this system by plotting $\dot{x}$ as a function of $x$, analogously to the cobweb diagram. This is demonstrated in Fig. 1.3. At $x$ values, where $\dot{x}$ is positive $x(t)$ is growing, while at negative values it is decreasing. This means, that fixed points, at which $x(t) = \text{const.}$ correspond to intersections of the graphs with the horizontal axis.

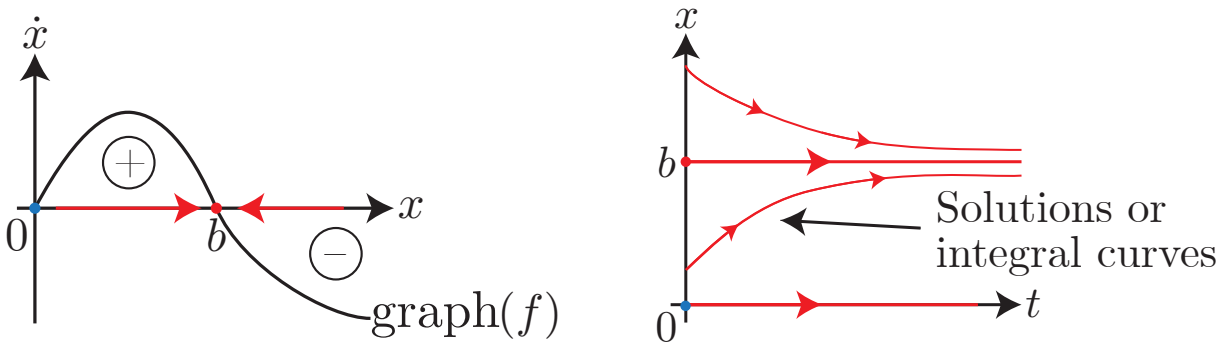


Figure 1.3: Left: Analysis of the right hand side. Right: Evolution in the extended phase space $P \times \mathbb{R}$.

*Example* 1.3 (Pendulum). Given the equation of motion

$$ml^2\ddot{\varphi} = -mgl\sin(\varphi). \tag{1.10}$$

We let $x_1 = \varphi$ and $x_2 = \dot{\varphi}$ to transform into the first-order ODE form

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = -\dfrac{g}{l}\sin(x_1). \end{cases} \tag{1.11}$$

Thus we have

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}; \quad f(x) = \begin{pmatrix} x_2 \\ -\frac{g}{l}\sin(x_1) \end{pmatrix}. \tag{1.12}$$

Qualitative analysis gives the following facts

- $(x_1, x_2) = (0, 0)$ and $(x_1, x_2) = (\pi, 0)$ are zeros of $f$.
- Energy is conserved, hence both small and large amplitude oscillations are expected.
- The function $f(x)$ has symmetries: it is invariant under the transformation $(x_1, x_2, t) \mapsto (x_1, -x_2, -t)$ and $(x_1, x_2, t) \mapsto (-x_1, x_2, -t)$. See the left panel of Fig. 1.4.

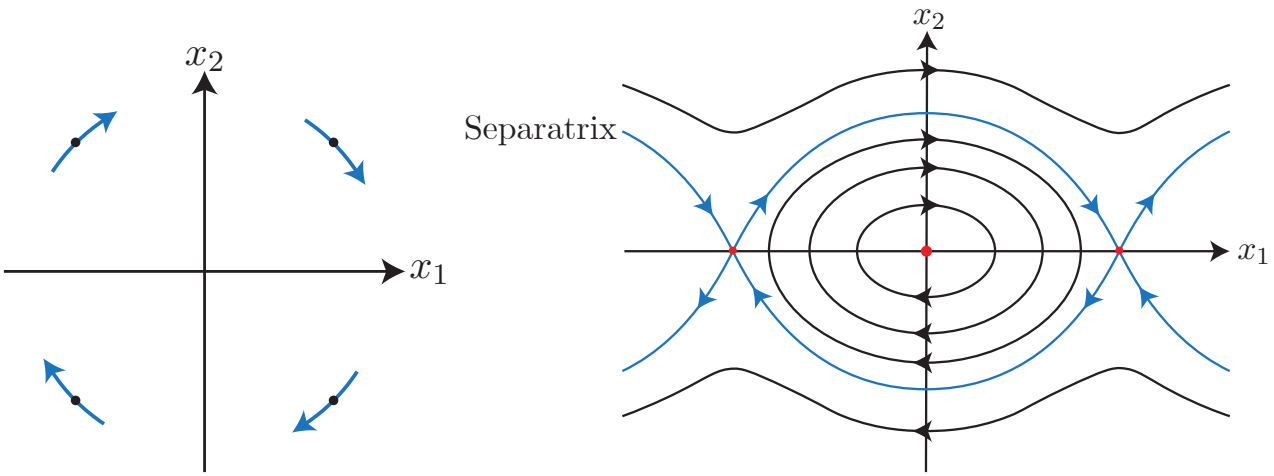


Figure 1.4: Left: The symmetries of the dynamical system. Right: Phase portrait of the pendulum. Red dots show the fixed points, while the blue trajectories make up the separatrix.

*Definition* 1.2. A separatrix is a boundary (i.e. a codimension-1 surface) in phase space which separates regions of qualitatively different behaviors. In practice, it is unobservable by itself and connects different fixed points. The separatrix of the pendulum is shown in the right panel of Fig. 4.

*Example* 1.4 (Exploit geometry of phase space for analysis). Consider two cities, $A$ and $B$. The two cities are connected by two roads, denoted by the blue and green curves of the left panel of Fig. 1.5. We assume that travelling on the two roads, it is possible for two bikes to make it from $A$ to $B$ without ever being further away from each other than a distance $d < D$.

Assume two trucks are trying to make it between $A$ and $B$, on different roads in the opposite direction, carrying a load of width $D$. Given this information, can the trucks make it without hitting each other? We can view this problem as a continuous dynamical system with two coordinates $x_1$ and $x_2$ that parameterize the two routes between $A$ and $B$. This dynamical system is, in general, non autonomous.

The right panel of Fig 1.5 shows the trajectories of the two trucks and the two bikes in phase space. The two trajectories must intersect by continuity, thus at that point the trucks

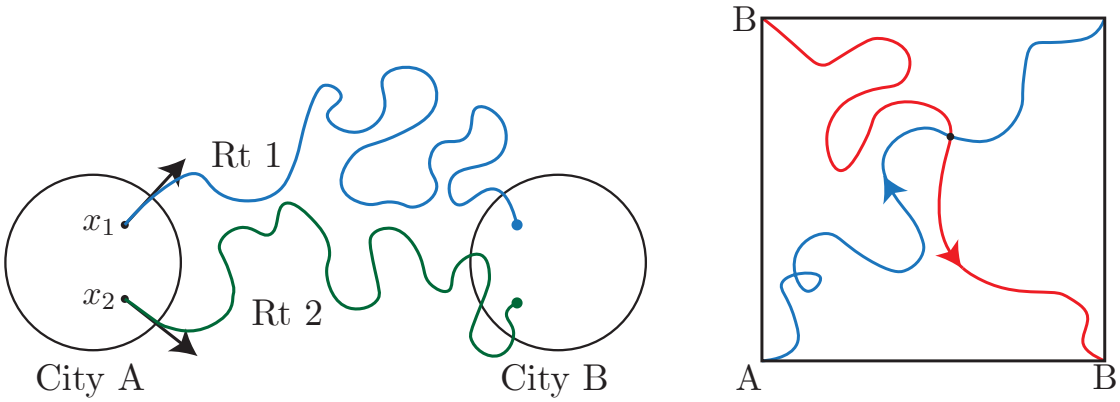


Figure 1.5: Left: An example of the two bike routes. Right: Blue represents the trajectory of the two bikes, red represents the trajectory of the two trucks.

must be at the same positions as the bikes, implying they are within distance $D$. Therefore the trucks must crash!

**Exercise 1.1.** *Consider a pendulum that strikes an inclined wall repeatedly, as shown in Fig. 1.6 below. Using the phase portrait of the pendulum discussed in class, sketch the trajectories in the phase space of this impact dynamical system for positive and negative values of the angle $\alpha$, when*

*(i) there us no loss of energy at impact*

*(ii) the coefficient of restitution is 0.5.*

*Identify the asymptotic behavior of the pendulum in each case.*

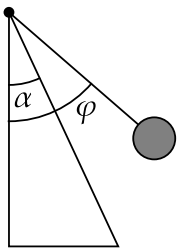


Figure 1.6: Pendulum attached to the inclined wall.

*Hint: During impact, the pre-impact velocity of the pendulum is assumed to jump instantaneously to its post-impact value, while the pendulum's position remains the same.*

# Chapter 2

# Fundamentals

In this chapter, we first review some fundamental properties of continuous dynamical systems that will be used heavily in later chapters. As we will see, these technical results are interesting in their own right. They can help in interpreting or cross-checking numerical results or physical models for self-consistency or accuracy.

## 2.1 Existence and uniqueness of solutions

Consider

$$\begin{cases} \dot{x} = f(x,t); \quad x \in \mathbb{R}^n \\ x(t_0) = x_0 \end{cases}. \tag{2.1}$$

Does this initial value problem have a unique solution? We have the following theorems to help us answer that question.

**Theorem 2.1** (Peano)**.** *If $f \in \mathcal{C}^0$ near $(x_0, t_0)$, then there exists a local solution $\varphi(t)$, i.e.,*

$$\dot{\varphi}(t) = f(\varphi(t), t), \varphi(t_0) = x_0; \ \forall t \in (t_0 - \varepsilon, t_0 + \varepsilon); \ 0 < \varepsilon \ll 1. \tag{2.2}$$

*Example* 2.1 (Free falling mass). Consider a point mass of mass $m$ at position $x$. The acceleration due to gravity is denoted by $g$. Measuring the potential energy from the reference point $x = x_0$, we have the total energy is conserved.

$$\frac{1}{2}m\dot{x}^2 = mg(x - x_0). \tag{2.3}$$

This implies that

$$\begin{cases} \dot{x} = \sqrt{2g(x - x_0)} \\ x(0) = x_0 \end{cases} \tag{2.4}$$

on the set $P = \{x \in \mathbb{R} : \ x \geq x_0\}$. Therefore we have that $f \in \mathcal{C}^0$ in phase space, so by Peano's theorem (cf. Theorem 2.1), there exists a local solution. A schematic diagram is shown in Fig. 2.1. The solution is actually $x(t) = x_0 + \frac{g}{2}(t - t_0)^2$, however $x(t) = x_0$ is also a solution to

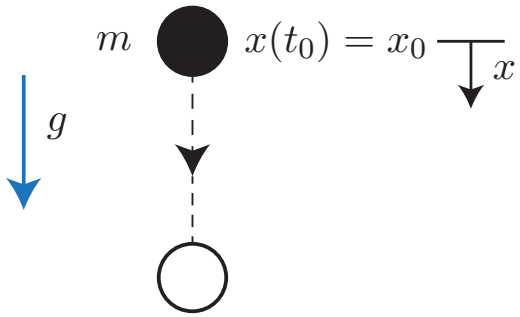


Figure 2.1: Schematic diagram of the point mass in free fall.

the IVP, therefore we do not have a unique solution. Physically there exists a solution, but this IVP was derived from a heuristic energy-principle, not from Newton's laws, which are not equivalent.

**Definition 2.1.** A function $f$ is called locally Lipschitz around $x_0$ if there exists an open set $U_{x_0}$ and $L > 0$ such that for all $x, y \in U_{x_0}$

$$\boxed{\|f(y,t) - f(x,t)\| \leq L\|y - x\|.} \tag{2.5}$$

*Example* 2.2 (Lipschitz functions). Fig. 2.2 shows an example of a Lipschitz and a non-Lipschitz function around $x_0$.

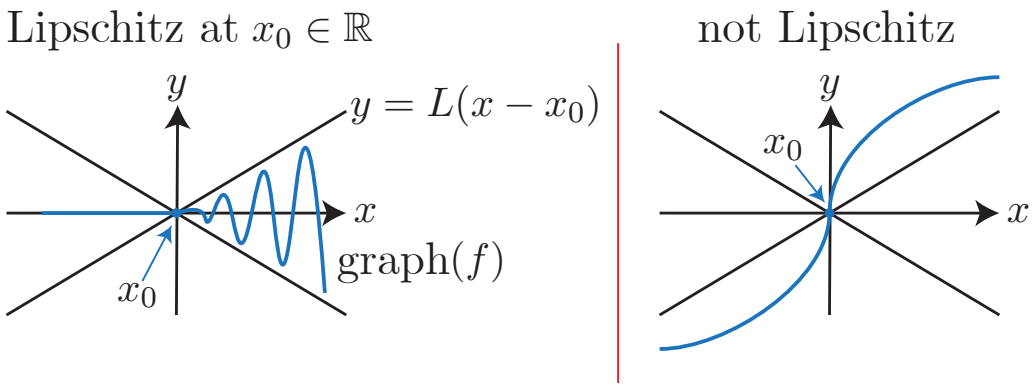


Figure 2.2: Interpretation of the Lipschitz property.

**Theorem 2.2** (Picard)**.** *Assume*

*(i)* $f \in \mathcal{C}^0$ *in* $t$ *near* $(t_0, x_0)$,

*(ii)* $f$ *is locally Lipschitz in* $x$ *near* $(t_0, x_0)$.

*Then there exists a unique local solution to the IVP. The proof can be found in [Arnold, 1992].*

**Note** the following relations. If $f$ is $\mathcal{C}^1 \implies f$ is Lipschitz $\implies f$ is $\mathcal{C}^0$.

*Example* 2.3 (Free falling mass revisted). We check if $f$ is Lipschitz.

$$\frac{|f(x) - f(x_0)|}{|x - x_0|} = \frac{\sqrt{2g}}{\sqrt{|x - x_0|}} \geq L. \tag{2.6}$$

Thus $f$ is not Lipschitz near $x_0$.

**Exercise 2.1.** *Many important properties of nonlinear dynamical systems follow from Gronwall's inequality. Assume that two positive, continuous scalar functions* $u(t)$ *and* $v(t)$ *satisfy the condition*

$$u(t) \leq C + \int_{t_0}^{t} u(\tau)v(\tau)\, d\tau \tag{2.7}$$

*for some constant* $C \geq 0$ *and for all* $t \geq t_0$. *Then Gronwall's inequality asserts that*

$$u(t) \leq Ce^{\int_{t_0}^{t} v(\tau)\, d\tau} \tag{2.8}$$

*for all* $t \geq t_0$. *The significance of this result is that it gives a* $u(t)$*-independent upper bound on the growth of* $u(t)$. *Using Gronwall's inequality, give an upper bound on how fast the solutions of a nonlinear ODE can separate from each other in time. In particular, show that for an ODE of the form*

$$\dot{x} = f(x, t); \quad x \in \mathbb{R}^n,$$

*and for two solutions starting from the initial conditions* $x_0$ *and* $\hat{x}_0$ *at time* $t_0$, *we have*

$$|x(t, x_0) - x(t, \hat{x}_0)| \leq |x_0 - \hat{x}_0|e^{L(t-t_0)},$$

*where* $L$ *is a Lipschitz constant for the function* $f$ *over a domain containing the trajectories of the system over the time interval* $[t_0, t]$.

*Hint: Substitute both solutions into the ODE, integrate the resulting two equations from* $t_0$ *to* $t$, *and estimate their normed difference.*

## 2.2 Geometric consequences of uniqueness

If the solution is unique, we have a few facts that can be derived from the geometric point of view.

(i) The trajectories of autonomous systems cannot intersect. Note that fixed points do not violate this (e.g. pendulum equations). See Fig. 2.3 which shows the phase portrait of the pendulum.

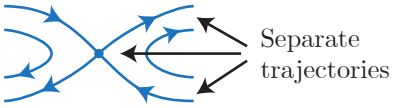


Figure 2.3: The phase portrait of the pendulum. Trajectories do not intersect since each arrow is pointing at separate trajectories.

(ii) For non-autonomous systems, intersections in phase space are possible: a trajectory may occupy the same point $x$ at a different time instants (see the left panel of Fig. 2.4. In this case we can extend the phase space in order to get an autonomous system where there cannot be any intersections.

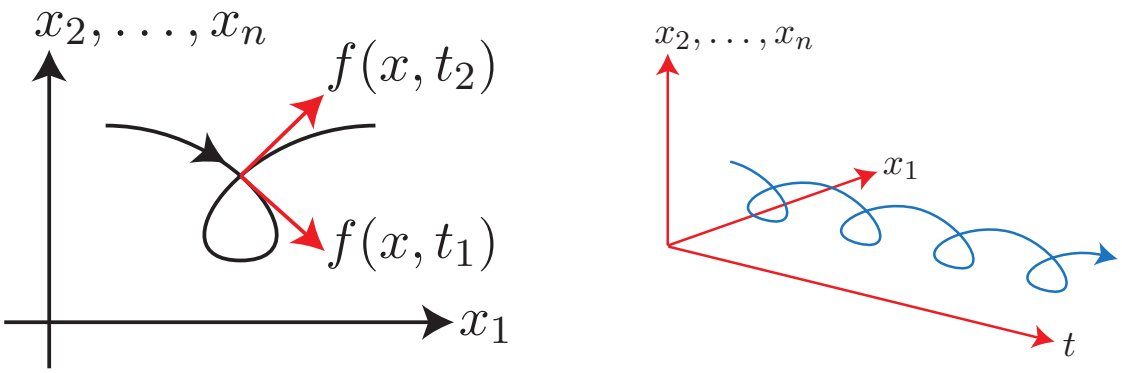


Figure 2.4: Left: Intersecting trajectories in phase space for a non-autonomous system. Right: The same trajectory in the extended phase space, without intersections.

$$X = \begin{pmatrix} x \\ t \end{pmatrix}, \; F(X) = \begin{pmatrix} f(x,t) \\ 1 \end{pmatrix}; \; \dot{X} = F(X). \tag{2.9}$$

## 2.3 Local vs global existence

*Example* 2.4 (Exploding solution).

$$\begin{cases} \dot{x} = x^2 \\ x(t_0) = 1. \end{cases} \tag{2.10}$$

Integrating yields the solution $x(t) = \frac{1}{1-(t-t_0)}$. This solution blows up at $t_\infty = t_0 + 1$, therefore the solution is only local. This is demonstrated in Fig. 2.5.

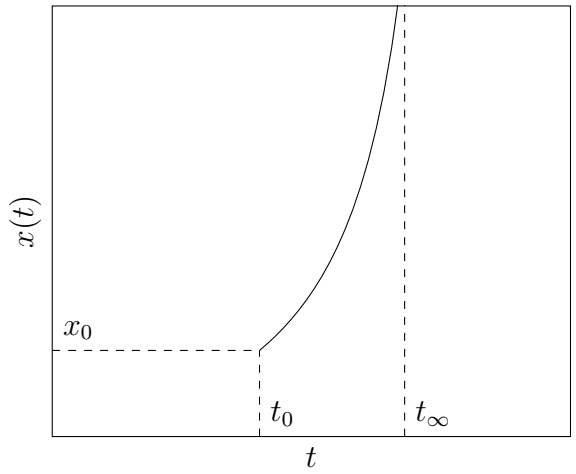


Figure 2.5: Solution to the ODE $\dot{x} = x^2$ started from $x(t_0) = 1$.

To address this problem of local solutions not being able to be continued into global solution, we have the following theorem.

**Theorem 2.3** (Continuation of solution)**.** *If a local solutions cannot be continued to a time $t = T$, then we must have*

$$\boxed{\lim_{t \to T} \|x(t)\| = \infty.} \tag{2.11}$$

*The proof can be found in [Arnold, 1992].*

*Example* 2.5 (Coupled Pendulum System). Consider two pendula of masses $m_1$ and $m_2$. They both have length $l$. The angles of these pendula are denoted by $\varphi_1$ and $\varphi_2$. Let us assume that they are coupled by a nonlinear spring, which can be described by a potential $V(\varphi_1, \varphi_2)$.

This setup is illustrated in Fig. 2.6. We set $x_1 = \varphi_1,\ x_2 = \dot{\varphi}_1,\ x_3 = \varphi_2,\ x_4 = \dot{\varphi}_2$ and get the following equation of motion

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = \dots \\ \dot{x}_3 = x_4 \\ \dot{x}_4 = \dots \end{cases} \tag{2.12}$$

The RHS is smooth, therefore there exists a unique local solution to any IVP. The phase space

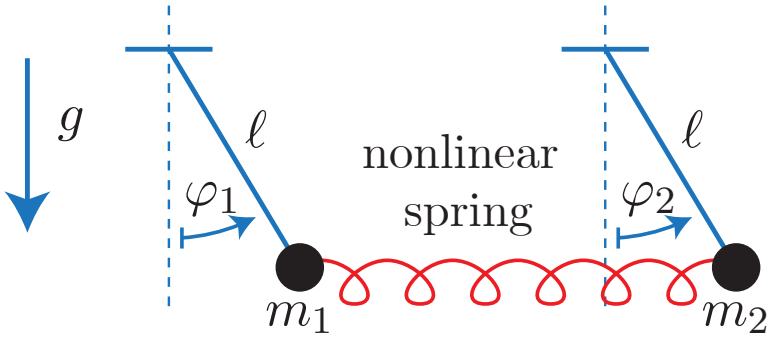


Figure 2.6: Physical setup of the coupled pendulum with a nonlinear spring.

is given by

$$P = \{x :\ x_1 \in S^1,\ x_2 \in \mathbb{R},\ x_3 \in S^1,\ x_4 \in \mathbb{R}\} = S^1 \times \mathbb{R} \times S^1 \times \mathbb{R}. \tag{2.13}$$

Where $S^1$ is the 1 dimensional sphere (i.e. a circle). With this space we know that $\|x_1\|$ and $\|x_3\|$ are bounded. Due to energy being conserved we have

$$E = T + V = \frac{1}{2} m_1 l_1 x_2^2 + \frac{1}{2} m_2 l_2 x_4^2 + \underbrace{V(x_1, x_3)}_{\geq 0} \tag{2.14a}$$

$$E = E_0 = \text{constant} \geq 0. \tag{2.14b}$$

Hence $\|x_2\|$ and $\|x_4\|$ are also bounded, therefore all solutions exist globally.

**Definition 2.2.** A linear system is one such that for $x \in \mathbb{R}^n,\ A(t) \in \mathbb{R}^{n \times n}$ and $A \in \mathcal{C}^0$

$$\boxed{\dot{x} = A(t)x.} \tag{2.15}$$

*Remark* 2.4. Note that $A$ can be written as $A = S + \Omega$ where $S = \frac{1}{2}(A + A^T)$ is symmetric (i.e. $S = S^T$) and $\Omega = \frac{1}{2}(A - A^T)$ is skew symmetric (i.e. $\Omega = -\Omega^T$). Furthermore the eigenvalues of $S$, $\lambda_i$, are all real and their respective eigenvectors, $e_i$, are orthogonal.

*Example* 2.6 (Global existence in linear systems).

$$\langle x, \dot{x}\rangle = \frac{1}{2}\frac{d}{dt}\|x(t)\|^2 = \langle x, A(t)x\rangle = \langle x, (S(t)+\Omega(t))x\rangle \tag{2.16a}$$

$$= \langle x, S(t)x\rangle + \underbrace{\langle x, \Omega(t)x\rangle}_{=0} \overset{(*)}{=} \sum_{i=1}^{n} \lambda_i(t)x_i^2 \tag{2.16b}$$

$$\leq \lambda_{\max}(t)\sum_{i=1}^{n} x_i^2 = \lambda_{\max}(t)\|x(t)\|^2. \tag{2.16c}$$

Where in $(*)$ we used that $x = \sum_{i=1}^{n} x_i e_i$ with $\|e_i\| = 1$ and $e_i \perp e_j$ for all $i \neq j$. Thus we get

$$\frac{\frac{1}{2}\frac{d}{dt}\|x(t)\|^2}{\|x(t)\|^2} \leq \lambda_{\max}(t) \implies \frac{1}{2}\log\left(\frac{\|x(t)\|^2}{\|x(t_0)\|^2}\right) \leq \int_{t_0}^{t} \lambda_{\max}(s)ds. \tag{2.17}$$

By exponentiating both sides, we obtain

$$\boxed{\|x(t)\| \leq \|x(t_0)\| \exp\left(\int_{t_0}^{t} \lambda_{\max}(s)ds\right).} \tag{2.18}$$

Therefore, by the continuation theorem, global solutions exist as long as $\int_{t_0}^{t} \lambda_{\max}(s)ds < \infty$.

## 2.4 Dependence on initial conditions

Given the IVP

$$\begin{cases} \dot{x} = f(x,t) \\ x(t_0) = x_0. \end{cases} \tag{2.19}$$

With $x \in \mathbb{R}^n$ and $f \in \mathcal{C}^r$ for some $r \geq 1$, we have the solution $x(t; t_0, x_0)$.

The dependence of the solution on initial data is of interest to us. This is due to us wanting the solution to be robust with respect to errors and uncertainties in the initial data. To address this, we have Theorem 2.5.

**Theorem 2.5.** *If $f \in \mathcal{C}^r$ for $r \geq 1$ then $x(t; t_0, x_0)$ is $\mathcal{C}^r$ in $(t_0, x_0)$. Proof in [Arnold, 1992].*

The geometric meaning of this is that for $U \subset P \subset \mathbb{R}^n$ we have that $F_{t_0}^t(U)$ is a smooth deformation of $U$ (cf. Fig. 2.7). It turns out $\left(F_{t_0}^t\right)^{-1} = F_t^{t_0}$ is also $\mathcal{C}^r$, hence we have that $F_{t_0}^t$ is a diffeomorphism.

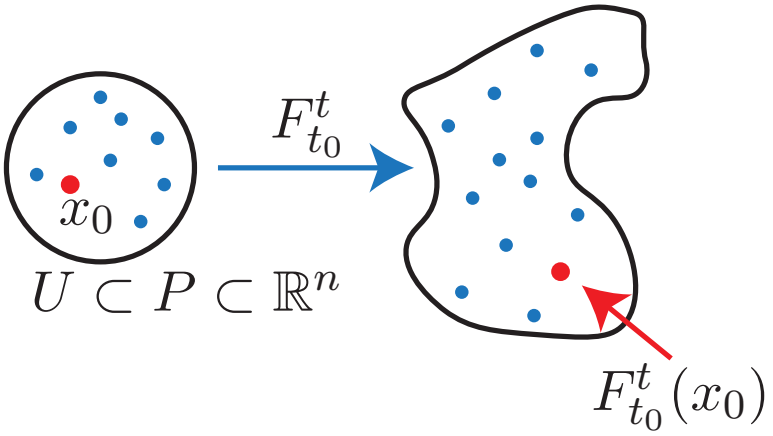


Figure 2.7: The smooth transformation of $U$. The red point on the right it $F^t_{t_0}(x_0)$, i.e. the image of $x_0$ under the evolution operator.

*Remark* 2.6 (The total differential). We denote the total differential of a function $f : \mathbb{R}^n \to \mathbb{R}^m$ as $Df$. The total differential is a function which takes a location $x$ as the argument and returns the derivative of $f$ at the point $x$, i.e. the Jacobian. This implies evaluating the Jacobian at the point $x$. For a function $f(x, y) = f(x_1, \ldots, x_n, y_1, \ldots, y_m) : \mathbb{R}^{n+m} \to \mathbb{R}^k$ the total differential $Df$ means the differential with respect to all of the variables and the total differential with respect to $x$, written $D_x f$ is the total differential only taken with respect to the $x$ variables. Thus for $f(x, y) : \mathbb{R}^{n+m} \to \mathbb{R}^k$ we have the total differential

$$Df = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} & \frac{\partial f_1}{\partial y_1} & \cdots & \frac{\partial f_1}{\partial y_m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial f_k}{\partial x_1} & \cdots & \frac{\partial f_k}{\partial x_n} & \frac{\partial f_k}{\partial y_1} & \cdots & \frac{\partial f_k}{\partial y_m} \end{pmatrix}; \tag{2.20}$$

$$Df(x_0, y_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x_0, y_0) & \cdots & \frac{\partial f_1}{\partial x_n}(x_0, y_0) & \frac{\partial f_1}{\partial y_1}(x_0, y_0) & \cdots & \frac{\partial f_1}{\partial y_m}(x_0, y_0) \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial f_k}{\partial x_1}(x_0, y_0) & \cdots & \frac{\partial f_k}{\partial x_n}(x_0, y_0) & \frac{\partial f_k}{\partial y_1}(x_0, y_0) & \cdots & \frac{\partial f_k}{\partial y_m}(x_0, y_0) \end{pmatrix}, \tag{2.21}$$

and the total differential with respect to $x$

$$D_x f = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_k}{\partial x_1} & \cdots & \frac{\partial f_k}{\partial x_n} \end{pmatrix}; \quad D_x f(x_0, y_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x_0, y_0) & \cdots & \frac{\partial f_1}{\partial x_n}(x_0, y_0) \\ \vdots & & \vdots \\ \frac{\partial f_k}{\partial x_1}(x_0, y_0) & \cdots & \frac{\partial f_k}{\partial x_n}(x_0, y_0) \end{pmatrix}. \tag{2.22}$$

Now, how can we compute the Jacobian of the flow map $\frac{\partial x(t;t_0,x_0)}{\partial x_0} = DF^t_{t_0}(x_0)$? We start from the IVP and take the gradient (with respect to $x_0$) of both sides. On the left hand side

we can exchange the order of the time derivative and the gradient and on the right hand side we use the chain rule. We end up with the equation

$$\frac{d}{dt}\frac{\partial x}{\partial x_0} = D_x f(x(t;t_0,x_0),t)\frac{\partial x}{\partial x_0}; \quad \frac{\partial x}{\partial x_0} \in \mathbb{R}^{n\times n}. \tag{2.23}$$

This means, that the flow map gradient satisfies the IVP

$$\frac{d}{dt}\left[DF_{t_0}^t(x_0)\right] = D_x f(F_{t_0}^t(x_0),t)DF_{t_0}^t(x_0) \tag{2.24a}$$

$$DF_{t_0}^{t_0}(x_0) = I. \tag{2.24b}$$

This is called the equation of variations, which is a linear, non-autonomous ODE for the matrix $M = DF_{t_0}^t(x_0)$

$$\begin{cases} \dot{M} = D_x f(x(t;t_0,x_0))M \\ M(t_0) = I. \end{cases} \tag{2.25}$$

*Example* 2.7 (Locations of extreme deformation in phase space). We define

$$\xi(t) := \tilde{x}(t) - x(t) = x(t;t_0,\tilde{x}_0) - x(t;t_0,x_0) \tag{2.26a}$$

$$= x(t;t_0,x_0) + \frac{\partial x}{\partial x_0}(t;t_0,x_0)\xi_0 + \mathcal{O}(\|\xi_0\|^2) - x(t;t_0,x_0) \tag{2.26b}$$

$$= DF_{t_0}^t(x_0)\xi_0 + \mathcal{O}(\|\xi_0\|^2). \tag{2.26c}$$

Where we used the Taylor expansion and assume the perturbation to $x_0$ is small, i.e. $\|\xi_0\| \ll 1$. Therefore we have

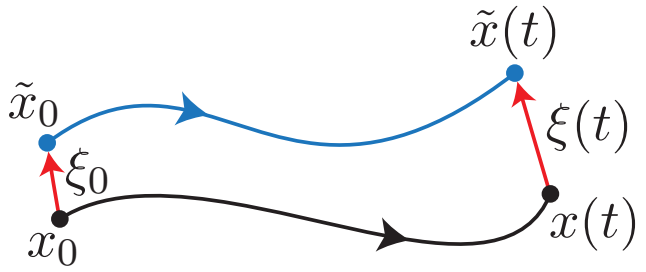


Figure 2.8: Illustration of the quantities $\xi_0$ and $\xi(t)$.

$$\|\xi(t)\|^2 = \langle DF_{t_0}^t(x_0)\xi_0, DF_{t_0}^t(x_0)\xi_0\rangle + \mathcal{O}(\|\xi_0\|^3) \tag{2.27a}$$

$$= \langle \xi_0, \underbrace{\left[DF_{t_0}^t(x_0)\right]^T DF_{t_0}^t(x_0)}_{=:C_{t_0}^t(x_0)}\xi_0\rangle + \mathcal{O}(\|\xi_0\|^3). \tag{2.27b}$$

$C_{t_0}^t(x_0)$ is known as the Cauchy-Green strain tensor. It is positive definite and symmetric and due to its dependence on the initial condition, $C_{t_0}^t(x_0)$ actually defines a *tensor field*. Therefore the largest possible deformation is

$$\max_{x_0,\xi_0} \frac{\|\xi(t)\|^2}{\|\xi_0\|^2} = \max_{x_0,\xi_0} \frac{\langle \xi_0, C_{t_0}^t(x_0)\xi_0\rangle}{\|\xi_0\|^2} = \max_{x_0} \lambda_n(x_0). \tag{2.28}$$

Where we used that $C_{t_0}^t$ is positive definite in the last equality, and that $\lambda_n(x_0)$ is the largest eigenvalue of $C_{t_0}^t(x_0)$. Because we typically have exponential growth we introduce the following quantity.

**Definition 2.3.** The finite-time Lyapunov exponent is defined as

$$\boxed{\mathrm{FTLE}_{t_0}^t(x_0) := \frac{1}{2|t-t_0|}\log(\lambda_n(x_0)).} \tag{2.29}$$

The FTLE is a diagnostic quantity for Lagrangian Coherent Structures (LCS), i.e. influential surfaces governing the evolution in $P$.

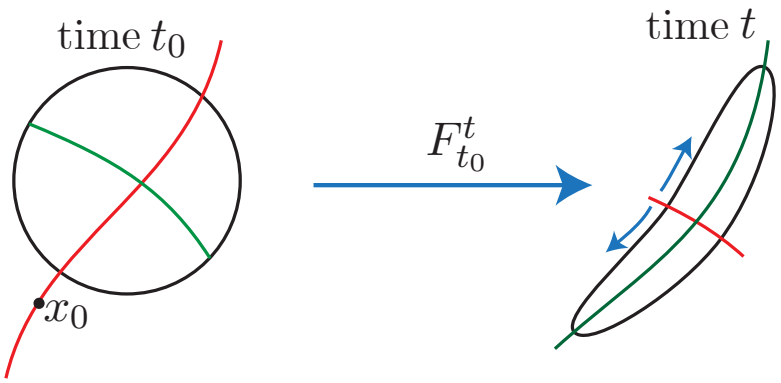


Figure 2.9: On the left the red ridge represents large values of $\mathrm{FTLE}_{t_0}^t$, on the right the green ridge the high values of $\mathrm{FTLE}_t^{t_0}$.

The ridges of $\mathrm{FTLE}_{t_0}^t$ are the repelling LCS, meanwhile the ridges of $\mathrm{FTLE}_t^{t_0}$ are the attracting LCS as depicted in Fig. 2.9. Now we are left with the problem of computing $F_{t_0}^t(x_0)$. Recall that analytically we start with $F_{t_0}^t(x_0)$ and use this to calculate $DF_{t_0}^t(x_0)$. From here we can find $C_{t_0}^t(x_0)$, giving us $\lambda_n(x_0)$ and thereby the FTLE. We now outline a process to compute the FTLE numerically.

(i) Define an initial $M \times N$ grid of initial data $x_0(i,j) \in \mathbb{R}^2$.

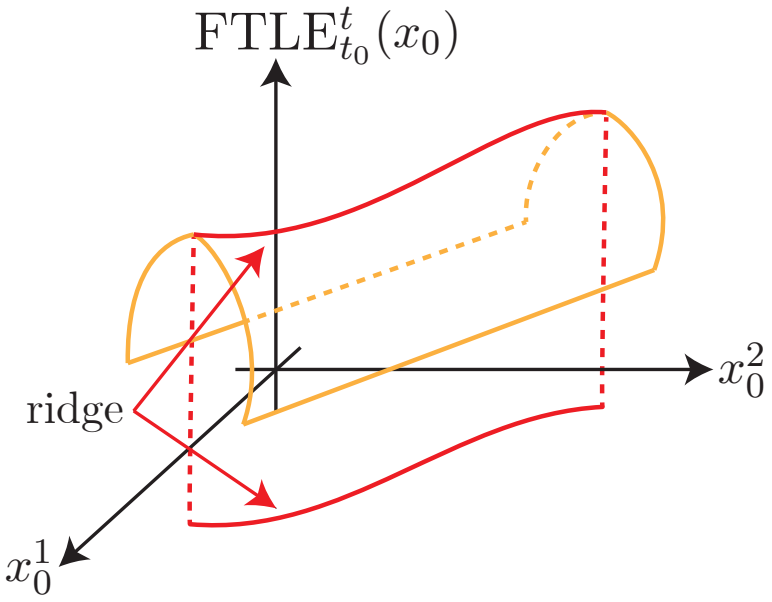


Figure 2.10: The projection of the FTLE ridge onto the initial value space.

(ii) Launch trajectories numerically from grid points to obtain a discrete approximation of $F_{t_0}^t(x_0)$ as $F_{t_0}^t(x_0(i,j))$.

(iii) Use finite differencing to approximate

$$DF_{t_0}^t(x_0(i,j)) \approx \begin{pmatrix} \frac{x(t;t_0,x_0(i,j)+\delta e_1)_1 - x(t;t_0,x_0(i,j)-\delta e_1)_1}{2\delta} & \cdots & \frac{x(t;t_0,x_0(i,j)+\delta e_n)_1 - x(t;t_0,x_0(i,j)-\delta e_n)_1}{2\delta} \\ \vdots & & \vdots \\ \frac{x(t;t_0,x_0(i,j)+\delta e_1)_n - x(t;t_0,x_0(i,j)-\delta e_1)_n}{2\delta} & \cdots & \frac{x(t;t_0,x_0(i,j)+\delta e_n)_n - x(t;t_0,x_0(i,j)-\delta e_n)_n}{2\delta} \end{pmatrix}. \tag{2.30}$$

This process then yields the surface we see in Fig. 2.10.

*Example* 2.8 (Calculating the FTLE for the double gyre). Due to incompressibility, we can define the two dimensional flow using a single scalar function called the stream function.

$$\Psi(x,y) = -\sin(\pi x)\sin(\pi y). \tag{2.31}$$

The components $(u,v)$ of the fluid velocity $(v = (u,v))$ are obtained as partial derivatives of the stream function, according to the formulas

$$\begin{cases} u = \dfrac{\partial \Psi}{\partial y} \\ v = -\dfrac{\partial \Psi}{\partial x}. \end{cases} \tag{2.32}$$

The Lagrangian trajectories of fluid particles obey the differential equations (i.e. we have the fluid velocity field)

$$\begin{cases} \dot{x} = u = \dfrac{\partial \Psi}{\partial y} \\ \dot{y} = v = -\dfrac{\partial \Psi}{\partial x}. \end{cases} \tag{2.33}$$

Interestingly, in this case, the phase space coincides with the physical space spanned by the coordinates $(x, y)$.

*Remark* 2.7. This is an example of a Hamiltonian system, where $\Psi$ is the Hamiltonian (usually denoted as $H$).

For any autonomous Hamiltonian system we have that the Hamiltonian is constant along trajectories. We can verify this as follows

$$\frac{d}{dt}\Psi(x(t), y(t)) = \frac{\partial \Psi}{\partial x}\dot{x} + \frac{\partial \Psi}{\partial y}\dot{y} = 0. \tag{2.34}$$

So we have that trajectories are level curves of $\Psi(x, y)$. We can then derive the phase portrait from the level curves of $\Psi$. Further, we have that $\dot{x} = \frac{\partial \Psi}{\partial y} = -\pi \sin(\pi x)\cos(\pi y)$ which yields that $\text{sign}(\dot{x}) = -\text{sign}(\sin(\pi x))\text{sign}(\cos(\pi y))$. Putting these together we can construct the contour plot with arrows. The contour plot, and FTLE approximation are shown in Fig. 2.11.

*Example* 2.9 (ABC flow). Let our dynamical system be defined as follows with $A, B, C \in \mathbb{R}$

$$\begin{cases} \dot{x} = A\sin(z) + C\cos(y) \\ \dot{y} = B\sin(x) + A\cos(z) \\ \dot{z} = C\sin(y) + B\cos(x). \end{cases} \tag{2.35}$$

This is an exact solution to Euler's equations. We have an autonomous velocity field. Depending on parameters it can even generate chaotic fluid trajectories. The numerical approximation of the FTLE for the ABC flow is depicted in Fig. 2.12.

**Exercise 2.2.** *Consider the non-dimensionalized, force-damped pendulum equation*

$$\ddot{x} + k\dot{x} + \sin(x) = a\sin(t), \tag{2.36}$$

*where $k \geq 0$ is the damping coefficient and $a \geq 0$ is the forcing amplitude.*

*(a) For vanishing damping and forcing ($a = k = 0$), compute and plot numerically the FTLE field for this system over a $100 \times 100$ grid of initial conditions, covering the square $[-\pi, \pi] \times [-\pi, \pi]$ in the phase space of the pendulum. Perform the computation for long enough times so that the FTLE plot fully reveals the separatrices of the system, as we discussed in class for the undamped pendulum.*

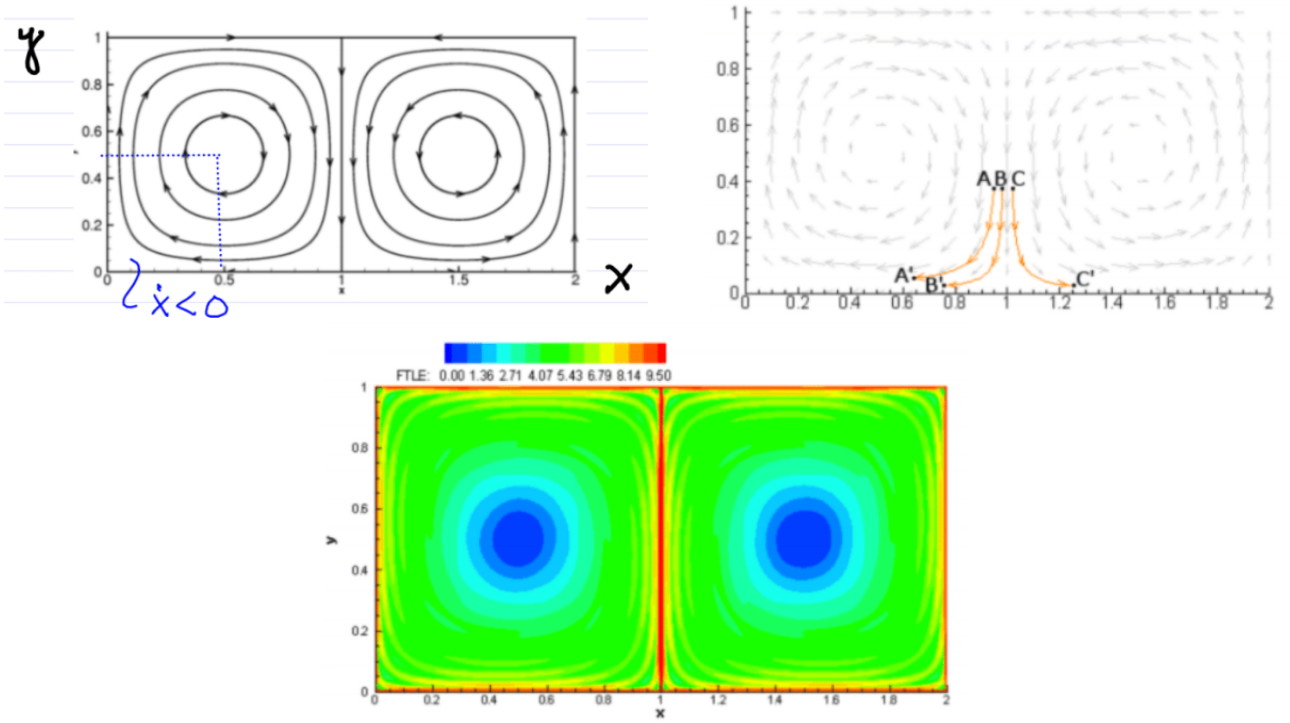


Figure 2.11: Top left: The analytic phase plot. Top right: The exploration done to calculate FTLE. Bottom: The FTLE plot. Figures here were taken from Shawn Shadden of UC Berkeley.

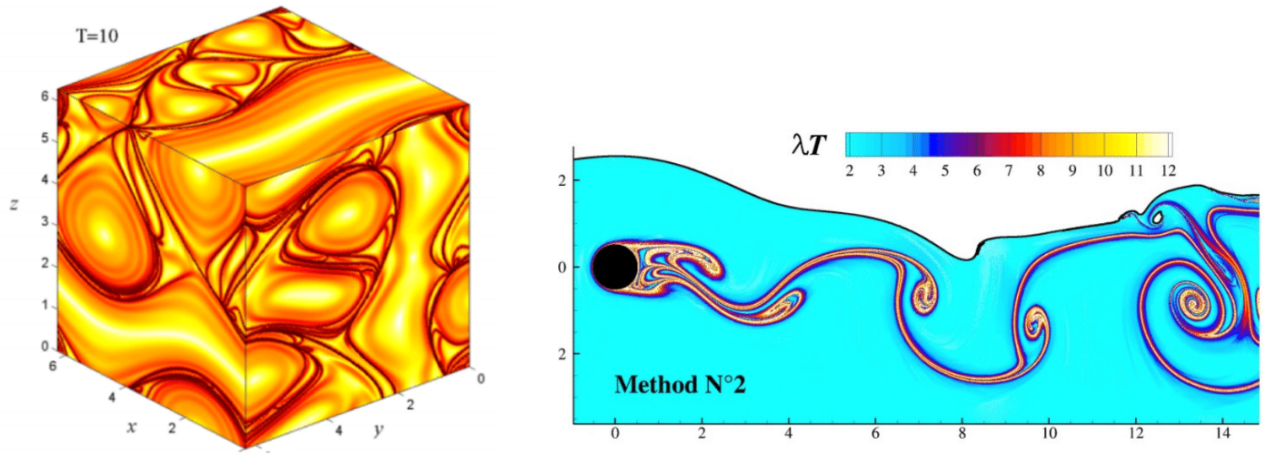


Figure 2.12: Left: numerically calculated FTLE field of the ABC flow. Darker colors signify higher FTLE values [Haller, 2001]. Right: Again the FTLE is plotted, for vortex shedding behind a cylinder under a free surface [Sun et al., 2016].

*(b) To explore the fate of these separatrices under mild damping and forcing, repeat the same FTLE computations for $a = 0.5$ and $k = 0.1$. Discuss domains of attractions and their boundaries based on the results.*

*Hint: Consider using MATLAB's ODE45 routine for trajectory integration. To this end, you need to assemble all initial grid points into a single long vector, and pass that vector on to ODE45 as an initial condition. You will then need to define an appropriate extended ODE, whose right-hand side will be called by ODE45 to produce an advected image of the extended initial condition vector. Subsequently, you will need to convert the elements of this advected vector back into advected grid positions. Alternatively, you could write your own 4-point Runge-Kutta solver to track trajectories simultaneously on the full grid of initial conditions (see the algorithm discussed in class).*

## 2.5 Dependence on parameters

We now have the IVP

$$\begin{cases} \dot{x} = f(x, t, \mu) \\ x(t_0) = x_0. \end{cases} \tag{2.37}$$

With $x \in \mathbb{R}^n,\ f \in \mathcal{C}^r,\ r \geq 1$, therefore we have a solution $x(t; t_0, x_0, \mu) \in \mathcal{C}^r_{x_0}$.

We now examine how solutions depend $\mu$. This is critical as solutions should be robust to changes or uncertainties in the model.

*Example* 2.10 (Perturbation Theory). Given a weakly nonlinear oscillator

$$m\ddot{x} + c\dot{x} + kx = \varepsilon f(x, \dot{x}, t),\ 0 \leq \varepsilon \ll 1,\ x \in \mathbb{R}. \tag{2.38}$$

The usual approach is to seek solutions by expanding from the known solution of the linear limit $\varepsilon = 0$, i.e.

$$x_\varepsilon(t) = \varphi_0(t) + \varepsilon\varphi_1(t) + \varepsilon^2\varphi_2(t) + \ldots + \mathcal{O}(\varepsilon^r). \tag{2.39}$$

If $x_\varepsilon(t)$ is in $\mathcal{C}^r_\varepsilon$, we have $\varphi_1(t) = \frac{\partial x_\varepsilon(t)}{\partial \varepsilon}\Big|_{\varepsilon=0}$ and $\varphi_2(t) = \frac{\partial^2 x_\varepsilon(t)}{\partial \varepsilon^2}\Big|_{\varepsilon=0}$

Regularity with respect to the parameter $\mu$ actually follows from regularity with respect to the initial condition $x_0$. We can use the following trick to extend the IVP with a dummy variable $\mu$

$$\begin{cases} \dot{x} = f(x, t, \mu) \\ \dot{\mu} = 0 \\ x(t_0) = x_0 \\ \mu(t_0) = \mu_0. \end{cases} \tag{2.40}$$

Thus with $X = \begin{pmatrix} x \\ \mu \end{pmatrix} \in \mathbb{R}^{n+p}$ and $F(X_0) = \begin{pmatrix} f \\ 0 \end{pmatrix}$; $X_0 = \begin{pmatrix} x_0 \\ \mu_0 \end{pmatrix}$. We have the extended IVP

$$\begin{cases} \dot{X} = F(X) \\ X(t_0) = X_0. \end{cases} \tag{2.41}$$

Applying the previous result on regularity with respect to $x_0$ to (2.41), we have that $f \in \mathcal{C}^r_{x,\mu}$ implies that $X(t) \in \mathcal{C}^r_{X_0}$ in turn implying that $x(t; t_0, x_0, \cdot) \in \mathcal{C}^r_\mu$. The solution is as smooth in parameters as the RHS of the dynamical system.

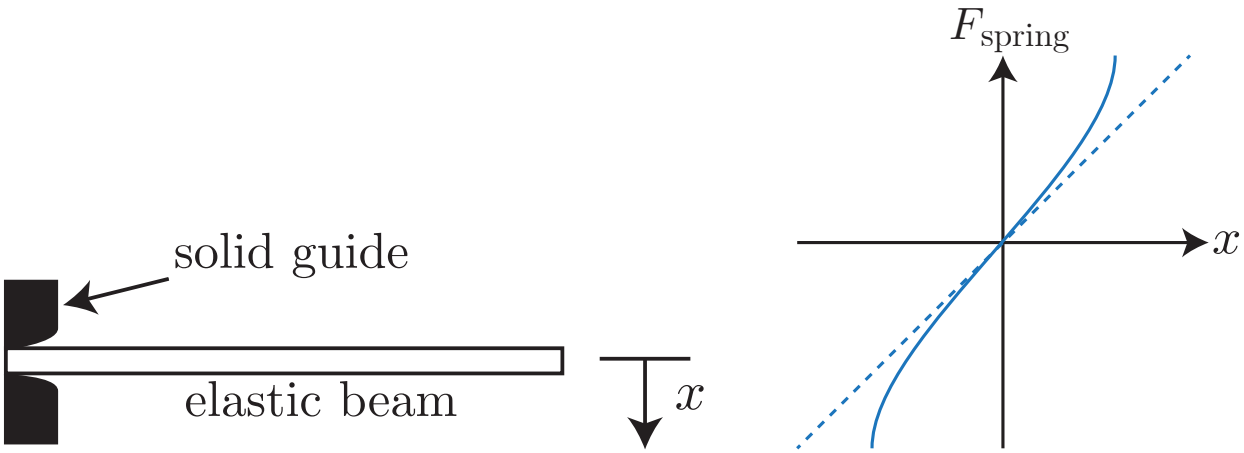


Figure 2.13: Setup for the nonlinear springboard.

*Example* 2.11 (Periodic Oscillations of a nonlinear springboard). Given an elastic beam extending from a solid guide, we measure the deflection of this beam with the variable $x$. This system is illustrated in the left panel of Fig. 2.13. By increasing $x$, the effective free length of the beam is shortened, thereby stiffening the spring nonlinearly. The effect of this nonlinearity on the force exerted on the spring is illustrated in the right panel of Fig. 2.13. This setup yields the following equations of motion

$$\begin{cases} \ddot{x} + x + \varepsilon x^3 = 0; \quad 0 \leq \varepsilon \ll 1 \\ x(0) = a_0; \quad \dot{x}(0) = 0. \end{cases} \tag{2.42}$$

So we have weak nonlinearity with no known explicit solution. Although weak, this nonlinearity is still significant, as can be seen in Fig. 2.14. Rewriting this as a first order ODE ($x_1 = x$; $x_2 = \dot{x}$ ), and note that the RHS is $\mathcal{C}^r_{x,\mu}$, therefore there exists a unique local solution that is also $\mathcal{C}^r_\mu$. Thus the expansion is justified

$$x_\varepsilon(t) = \varphi_0(t) + \varepsilon\varphi_1(t) + \ldots + \mathcal{O}(\varepsilon^r). \tag{2.43}$$

We can see, by substitution, that for $\varepsilon = 0$ we find that $\varphi_0(t) = a_0 \cos(t)$.

Now we look specifically for $T$-periodic solutions, as we would expect such a solution physically, therefore we have

$$\varphi_i(t) = \varphi_i(t+T). \tag{2.44}$$

The period $T$ still has to be determined. Plugging this power series into (2.43) to get

$$\mathcal{O}(1): \quad \ddot{\varphi}_0 + \varphi_0 = 0 \tag{2.45a}$$

$$\mathcal{O}(\varepsilon): \quad \ddot{\varphi}_1 + \underbrace{\varphi_1}_{\omega=1} = -\varphi_0^3 = -a_0^3 \cos^3(t) = -a_0^3 \left[ \frac{1}{4}\cos(3t) + \frac{3}{4} \underbrace{\cos(t)}_{\text{resonance}} \right]. \tag{2.45b}$$

We can see that (2.45b) is a linear oscillator with a forcing coming from the zeroth order solution. Since the zeroth order solution $\varphi_0 = a_0 \cos(t)$ already solves the IVP we have the following initial conditions

$$\varphi_1(0) = 0; \quad \dot{\varphi}_1(t) = 0. \tag{2.46}$$

This holds as $\varphi_0 = a_0 \cos(t)$ already solves the IVP. The general solution to this equation is the sum of two terms. We add the general solution of the homogeneous part and a particular solution to the inhomogeneous part. We can write this solution to (2.45b) as

$$\varphi_1(t) = \varphi_1^{\text{hom}}(t) + \varphi_1^{\text{part}}(t) \tag{2.47a}$$

$$= \underbrace{A\cos(t) + B\sin(t)}_{\text{TBD from initial conditions}} + \underbrace{C\cos(3t) + Dt\cos(t) + Et\sin(t)}_{\text{TBD from (2.45b)}}. \tag{2.47b}$$

Observe that due to a resonance between the natural frequency of the oscillator and the forcing secular terms, $t\cos(t)$ and $t\sin(t)$ appear. Thus it cannot be periodic, so our Ansatz already fails for $i = 1$. We conclude that no solution of this type exists. Our Ansatz was too restrictive and $T$ should depend on $\varepsilon$.

**Lindstedt's idea** We should seek a solution of the form

$$x_\varepsilon(t) = \varphi_0(t;\varepsilon) + \varepsilon\varphi_1(t;\varepsilon) + \varepsilon^2\varphi_2(t;\varepsilon) + \mathcal{O}(\varepsilon^3). \tag{2.48}$$

Furthermore $\varphi_i$ should be $T_\varepsilon$ periodic, i.e. the period should depend on the strength of the nonlinearity $\varepsilon$.

$$\varphi_i(t+T_\varepsilon;\varepsilon) = \varphi_i(t;\varepsilon). \tag{2.49}$$

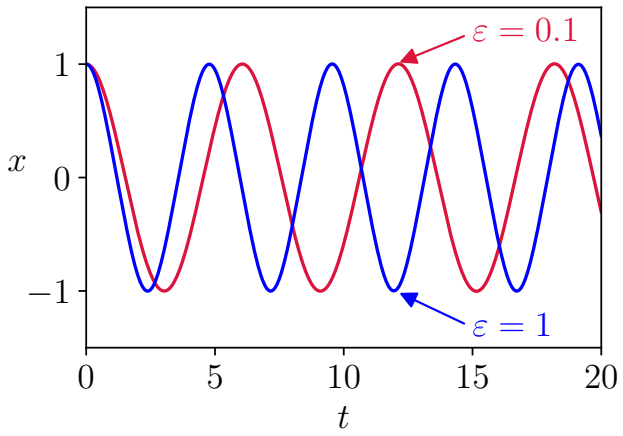


Figure 2.14: Numerical integration of $x$ for $a_0 = 1$ and different values of $\varepsilon$.

Rewriting the period as

$$T_\varepsilon = \frac{2\pi}{\omega(\varepsilon)}; \quad \omega(\varepsilon) = 1 + \varepsilon\omega_1 + \varepsilon^2\omega_2 + \mathcal{O}(\varepsilon^3). \tag{2.50}$$

We then rescale time according to $\tau = \omega(\varepsilon)t$ to find

$$\frac{d}{d\tau} = \frac{1}{\omega(\varepsilon)}\frac{d}{dt} \implies \boxed{[\omega(\varepsilon)]^2 x'' + x + \varepsilon x^3 = 0.} \tag{2.51}$$

Where we have taken $x'$ to represent $\frac{dx}{d\tau}$. Plugging our expression into the rescaled ODE yields

$$\left(1 + 2\varepsilon\omega_1 + \mathcal{O}(\varepsilon^2)\right)\left[\varphi_0'' + \varepsilon\varphi'' + \mathcal{O}(\varepsilon^2)\right] + \left[\varphi_0 + \varepsilon\varphi_1 + \mathcal{O}(\varepsilon^2)\right] + \varepsilon\left[\varphi_0^3 + \mathcal{O}(\varepsilon)\right] = 0. \tag{2.52}$$

Matching equal powers of $\varepsilon$ yields

$$\mathcal{O}(1): \ \varphi_0'' + \varphi_0 = 0 \implies \varphi_0(\tau) = a_0\cos(\tau); \quad \varphi_0(0) = a_0; \quad \dot\varphi_0(0) = 0 \tag{2.53a}$$

$$\mathcal{O}(\varepsilon): \ \varphi_1'' + \varphi_1 = -\varphi_0^3 - 2\omega_1\varphi_0'' = \left(2\omega_1 a_0 - \frac{3}{4}a_0^3\right)\underbrace{\cos(\tau)}_{\text{resonance}} - \frac{a_0^3}{4}\cos(3\tau); \tag{2.53b}$$

$$\varphi_1(0) = 0; \quad \dot\varphi_1(0) = 0. \tag{2.53c}$$

From the first line, we can see the initial conditions are fulfilled. In this step we used that $\dot\varphi(t=0) = 0$ if and only if $\omega(\varepsilon)\varphi'(0) = 0$. We get the solution

$$\varphi_1(t) = A\cos(\tau) + B\sin(\tau) + C\cos(3\tau) + D\tau\cos(\tau) + E\tau\sin(\tau). \tag{2.54}$$

The presence of resonance again excludes periodic solutions, but now we can select $\omega_1$ to eliminate these terms.

$$2\omega_1 a_0 - \frac{3}{4}a_0^3 = 0 \implies \boxed{\omega_1 = \frac{3}{8}a_0^2.} \tag{2.55}$$

This successfully eliminates the resonance and determines the missing frequency term at $\mathcal{O}(\varepsilon)$. Thus we find

$$x_\varepsilon(\tau) = a_0\cos(\tau) - \frac{\varepsilon}{32}a_0^3\left(\cos(\tau) - \cos(3\tau)\right) + \mathcal{O}(\varepsilon^2). \tag{2.56}$$

In the original time scaling this is

$$x_\varepsilon(t) = a_0\cos(\omega t) - \frac{\varepsilon}{32}a_0^3\left(\cos(\omega t) - \cos(3\omega t)\right) + \mathcal{O}(\varepsilon^2); \quad \omega = 1 + \frac{3}{8}\varepsilon a_0^2 + \mathcal{O}(\varepsilon^2). \tag{2.57}$$

This procedure can be continued to higher order terms, where we select $\omega_2$ so that the $\mathcal{O}(\varepsilon^2)$ terms cancel.

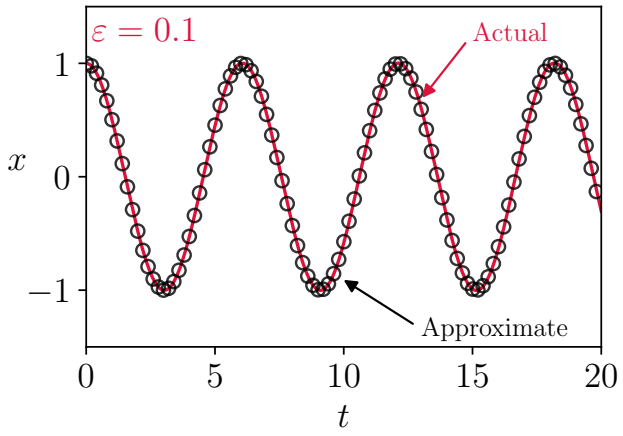


Figure 2.15: Approximation (dots) vs analytic solution (solid line) of $x$ on the time interval $[0, 20]$.

*Exercise* 2.3. *Consider the nonlinear oscillator*

$$\ddot{x} + \omega_0^2 x = \varepsilon M x^2, \tag{2.58}$$

*where $\varepsilon M x^2$ represents a small nonlinear forcing term ($0 \leq \varepsilon \ll 1, M > 0$).*

*Using Lindstedt's method, find an $\mathcal{O}(\varepsilon)$ approximation for nonlinear motions as a function of their initial position, with zero initial velocity.*

*Exercise* 2.4. *Consider the forced* van der Pol equation

$$\ddot{x} + \varepsilon(x^2 - 1)\dot{x} + x = F\cos(\omega t), \tag{2.59}$$

*which arises in models of self-excited oscillation, such as those of a valve generator with a cubic valve characteristic. Here* $F, \omega > 0$ *are parameters, and* $0 \leq \varepsilon \ll 1$.

*(i) For small values of* $\varepsilon$, *find an approximation for an* ***exactly*** $2\pi/\omega$*-periodic solution of the equation. The error of your approximation should be* $\mathcal{O}(\varepsilon)$.

*(ii) For* $\varepsilon = 0.1$, $\omega = 2$, *and* $F = 1$, *verify your prediction numerically by solving the equation numerically. Plot your numerical solution along with your analytic prediction computed in (i).*

*Note: For chaotic dynamics in the forced van der Pol equation, see Section 2.1 of [Guckenheimer and Holmes, 1990].*

# Chapter 3

# Stability of fixed points

Now we would like to begin to explore the behaviour of dynamical systems around fixed points. This will allow us to find out if we should expect to observe a fixed state, and to understand what happens if we perturb the system away from this fixed state.

## 3.1 Basic definitions

Consider

$$\dot{x} = f(x,t),\ x \in \mathbb{R}^n,\ f \in \mathcal{C}^1. \tag{3.1}$$

Assume that $x = 0$ is a fixed point, i.e. $f(0,t) = 0$ for all $t \in \mathbb{R}$. If the fixed point is originally at $x_0 \neq 0$, shift it to zero by letting $\tilde{x} := x - x_0$, therefore

$$\dot{\tilde{x}} = \dot{x} = f(x_0 + \tilde{x}, t) = \tilde{f}(\tilde{x}, t). \tag{3.2}$$

We would like to understand how the dynamical system behaves near its equilibrium state. To this end we introduce the following definitions.

**Definition 3.1** (Lyapupnov Stability)**.** The fixed point $x = 0$ is stable if for all $t_0$, for all $\varepsilon > 0$ small enough, there exists a $\delta = \delta(t_0, \varepsilon)$, such that for all $x_0 \in \mathbb{R}^n$ with $\|x_0\| \leq \delta$, we have

$$\boxed{\|x(t; t_0, x_0)\| \leq \varepsilon \quad \forall t \geq t_0.} \tag{3.3}$$

*Remark* 3.1 (N-dimensional ball). When writing $\mathcal{B}(r)$ we refer to the ball of radius $r$ in $\mathbb{R}^n$, i.e. the set $\{x :\ \|x\| < r\}$.

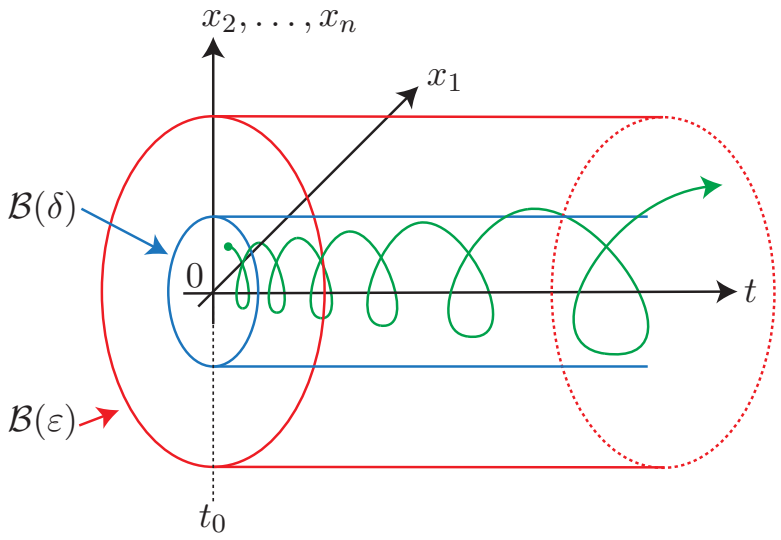


Figure 3.1: An example of $\delta$ for a given Lyapunov stable fixed point.

*Example* 3.1 (Stability of the lower equilibrium of the pendulum). Recall the equation of motion of the pendulum $\ddot{\varphi} + \sin(\varphi) = 0$, that we transform into a first order ODE by setting $x_1 = \varphi$ and $x_2 = \dot{\varphi}$ to obtain

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = -\sin(x_1). \end{cases} \tag{3.4}$$

For small $\varepsilon > 0$, this geometric procedure gives a $\delta(\varepsilon) > 0$ such that the definition of stability is satisfied for $x = 0$. We can see in Fig. 3.2 that for any initial point chosen within the blue circle, it's trajectory remains within the red circle for all time (cf. Fig. 3.1). Therefore $x = 0$ is (Lyapunov) stable.

**Definition 3.2** (Asymptotic stability)**.** The fixed point $x = 0$ is *asymptotically stable* if

(i) it is stable,

(ii) for all $t_0$, there exists $\delta_0(t_0, \varepsilon)$ such that for every $x_0$ with $\|x_0\| \leq \delta_0$ we have

$$\boxed{\lim_{t\to\infty} x(t; t_0, x_0) = 0.} \tag{3.5}$$

**Definition 3.3** (Domain of attraction)**.** The *domain of attraction* of the fixed point $x = 0$ is the set of all $x_0$'s for which

$$\boxed{\lim_{t\to\infty} x(t; t_0, x_0) = 0.} \tag{3.6}$$

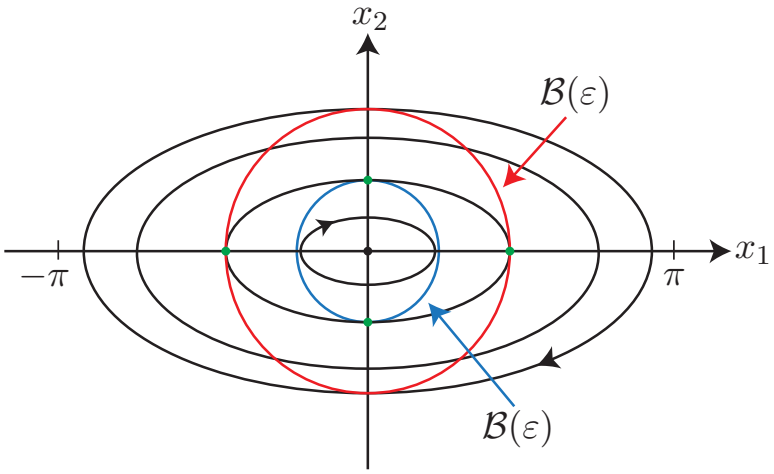


Figure 3.2: Stability of the lower equilibrium for the pendulum, here $0 < \varepsilon < \pi$.

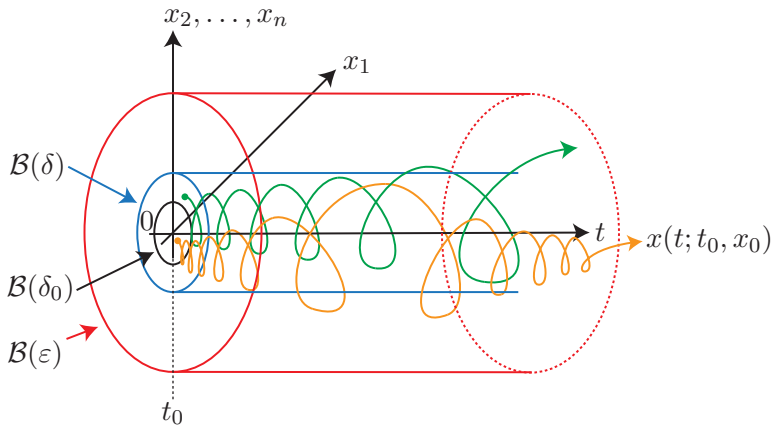


Figure 3.3: An example for an asymptotically stable fixed point (black trajectory).

*Example* 3.2 (Damped pendulum). We have the equation of motion with the linear damping coefficient $c$

$$\ddot{\varphi} + c\dot{\varphi} + \sin(\varphi) = 0, \quad c > 0. \tag{3.7}$$

Transforming into a first-order ODE with $x_1 = \varphi$ and $x_2 = \dot{\varphi}$ gives

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = -cx_2 - \sin(x_1). \end{cases} \tag{3.8}$$

The total energy is given by

$$E = \frac{1}{2}x_2^2 + (1 - \cos(x_1)). \tag{3.9}$$

Further we have the rate of energy change

$$\frac{d}{dt}E(x_1(t), x_2(t)) = x_2\left(\dot{x}_2 + \sin(x_1)\right) = -cx_2^2. \tag{3.10}$$

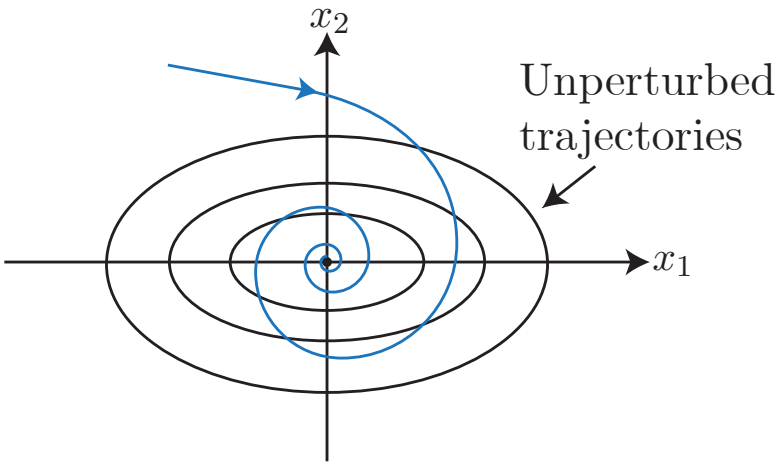


Figure 3.4: An example of a trajectory which loses energy, in this case due to damping.

Therefore, along trajectories energy decreases monotonically as shown in Fig 3.4. By the $\mathcal{C}^0$ dependence of the trajectory on initial conditions, the trajectories remain close to the undamped oscillations for small $c > 0$. We conclude that trajectories are inward spirals for a small dissapation $c > 0$. The fixed point $x = 0$ is still Lyapunov stable, but asymptotic stability does not yet follow (is the limit of $x(t)$ equal to 0?).

*Remark* 3.2 (LaSalle's invariance principle). This conclusion follows rigorously from LaSalle's invariance principle, namely if we assume that $\dot{x} = f(x)$, $f \in \mathcal{C}^1$, and that there exists a $V \in \mathcal{C}^1$ with

$$\dot{V} = \frac{dV(x(t))}{dt} \leq 0. \tag{3.11}$$

Then the set of accumulation points for any trajectory is contained in the set of trajectories that stay within the set $I = \{x \in \mathbb{R}^n : \dot{V}(x) = 0\}$.

*Example* 3.3. Consider the following dynamical system in polar coordinates, i.e. $r\cos(\theta) = x$ and $r\sin(\theta) = y$,

$$\begin{cases} \dot{r} = r(1-r) \\ \dot{\theta} = \sin^2\left(\dfrac{\theta}{2}\right). \end{cases} \tag{3.12}$$

Note that $r = 0$ is a fixed point, the set $r = 1$ is an invariant circle, and the set $\theta = 0$ is an invariant set. An invariant set is a set such that if the dynamical system is started on the set, it remains in the set for all time. Examining the radial evolution reveals that the equation of motion decouples. We see that $\dot{\theta} \geq 0$, so rotation is either positive or null.

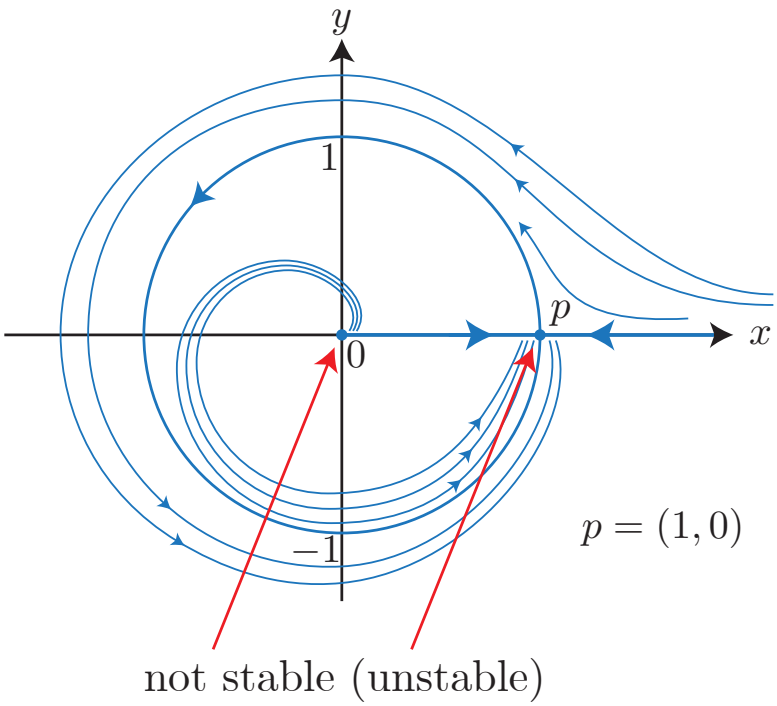


Figure 3.5: Phase portrait of the dynamical system in cartesian coordinates, with the red arrows pointing to the two unstable equilibria.

From Fig. 3.5 we can see that both of the fixed points, $(0, 0)$ and $(1, 0)$, are not stable. However, inspecting Fig. 3.5 we see that that $p = (1, 0)$ is an example of an attractor: a set with an open neighborhood of points that all approach the set as $t \to \infty$.

**Definition 3.4** (Invariant set)**.** The set $S \subset P$ is an *invariant set* for the flow map $F^t : P \to P$ if $F^t(S) = S$ for all $t \in \mathbb{R}$.

**Definition 3.5** (Unstable point)**.** A fixed point $x = 0$ is unstable if it is not stable.

*Remark* 3.3. We can negate a mathematical statement by using the reverse relational operators outside the statements involving these operators i.e. $\exists \to \forall$ and $\forall \to \exists$. For example we have for continuity $\forall\varepsilon\ \exists\delta :\ \|f(x) - f(y)\| < \varepsilon$ if $\|x - y\| < \delta$, meanwhile for discontinuity we have $\exists\varepsilon :\ \forall\delta :\ \|f(x) - f(y)\| \geq \varepsilon$ for $\|x - y\| < \delta$.

In our case for stability we have

$$\forall \varepsilon, t_0 : \quad \exists \delta > 0 : \quad \forall x_0 \text{ with } \|x_0\| < \delta : \quad \|x(t)\| \leq \varepsilon \quad \forall t \geq t_0. \tag{3.13}$$

Meanwhile for instability

$$\exists \varepsilon, t_0 : \quad \underbrace{\forall \delta > 0}_{\text{“for arbitrarily small”}} \quad : \quad \exists x_0 \text{ with } \|x_0\| < \delta : \quad \|x(t)\| > \varepsilon \quad \underbrace{\exists t \geq t_0}_{\text{“for some”}} . \tag{3.14}$$

This negation is demonstration in Fig. 3.6.

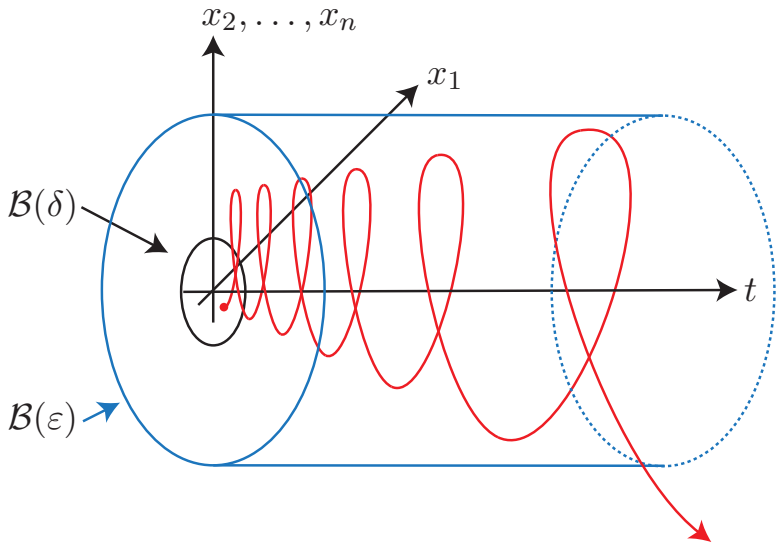


Figure 3.6: Example of an unstable fixed point, with the red trajectory representing a trajectory starting arbitrarily close to the fixed point, leaving a given $\varepsilon$-ball.

*Remark* 3.4. By $\mathcal{C}^0$ dependence of trajectories on initial conditions, if $x(t; t_0, x_0)$ leaves $\mathcal{B}(\varepsilon)$, then for $\tilde{x}_0$ close enough to $x_0$, $x(t; t_0, \tilde{x}_0)$ also leaves $\mathcal{B}(\varepsilon)$. Since this is true on an open set around $x_0$, the measure of such trajectories in nonzero, the instability is observable!

*Example* 3.4 (Unstable fixed point of pendulum). In contrast, we can have that infinitely many trajectories converge to the fixed point, yet it is still unstable, as illustrated in Fig. 3.7. In fact, the converging trajectories form a measure-zero set, thus the stability near the unstable equilibrium is unobservable.

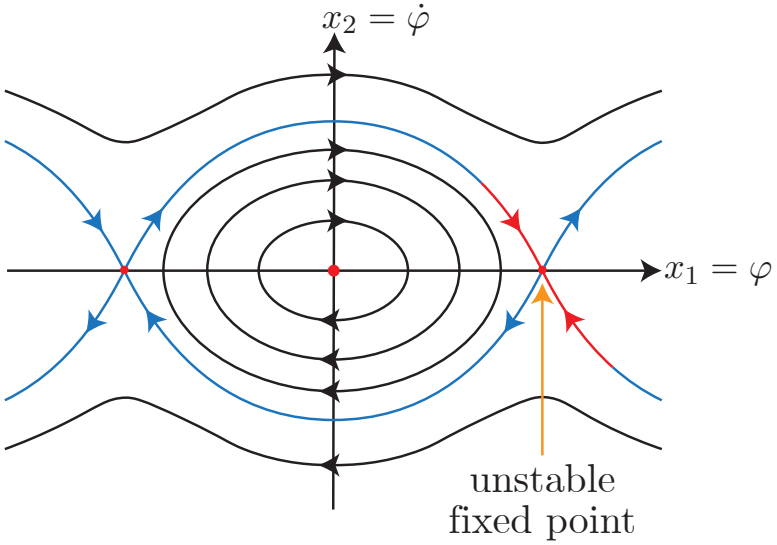


Figure 3.7: The phase portrait around the unstable fixed point of the pendulum, with the stable trajectories (red).

## 3.2 Stability based on linearization

We would like to derive a more general method to analyze the stability of fixed points, thus we try to simplify our system around the fixed point and discover what this can tell us about the full (unsimplified) system. In the following section we shall always assume that our system is autonomous. We will have the following setup

$$\dot{x} = f(x), \quad f \in \mathcal{C}^1, \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n, \quad p = \begin{pmatrix} p_1 \\ \vdots \\ p_n \end{pmatrix} \in \mathbb{R}^n. \tag{3.15}$$

If $f(p) = 0$, then $p$ is a fixed point. By transforming using $y = x - p$, we have that in the transformed system $y = 0$ is a fixed point. Furthermore, we have that around $y = 0$ the ODE is

$$\dot{y} = f(p + y) = \underbrace{f(p)}_{=0} + Df(p)y + o(\|y\|) = Df(p)y + o(\|y\|). \tag{3.16}$$

*Remark* 3.5. Since we only assumed one continuous derivative for the function $f$ the remainder term in the Taylor approximation is $o(\|y\|)$. The little o notation means that

$$\lim_{\|y\| \to 0} \frac{o(\|y\|)}{\|y\|} = 0. \tag{3.17}$$

**Definition 3.6** (Linearized ODE)**.** We define the *linearization* of (3.15) at the fixed point $p$ as

$$\boxed{\dot{y} = Ay; \quad y \in \mathbb{R}^n, \quad A := Df(p) \in \mathbb{R}_{n\times n}; \quad Df(p) = \left.\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix}\right|_{x=p}.} \tag{3.18}$$

Now we would like to study the stability of the fixed point $y = 0$ in (3.18). From this analysis, we want to know the relevance of our results for the full nonlinear system (3.15).

## 3.3 Review of linear dynamical systems

Recall the setup

$$\dot{y} = A(t)y, \quad y \in \mathbb{R}^n, \quad A \in \mathbb{R}^{n\times n}, \quad A \in \mathcal{C}^0_t. \tag{3.19}$$

The following facts have already been established

- We know that the global existence and uniqueness of solutions is guaranteed.
- The superposition principle holds; namely the linear combination of solutions is also a solution.
- There exists a set of $n$ linearly independent solutions: $\varphi_1(t), \ldots, \varphi_n(t) \in \mathbb{R}_n$.
- The general solution is

$$y(t) = \sum_{i=1}^{n} c_i\varphi_i(t) = \underbrace{\begin{bmatrix}\varphi_1(t) & \ldots & \varphi_n(t)\end{bmatrix}}_{\Psi(t)\text{: fundamental matrix solution}} \underbrace{\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}}_{c} = \Psi(t)c; \quad \dot{\Psi} = A(t)\Psi. \tag{3.20}$$

- We have the initial value problem $y(t_0) = y_0$ which implies

$$\Psi(t_0)c = y_0 \implies y(t) = \underbrace{\Psi(t)\left[\Psi(t_0)\right]^{-1}}_{\phi(t):=F^t_{t_0}} y_0 \tag{3.21}$$

  Where we used that the $\varphi_i(t)$ are linearly independent in the last equality. And we have the *normalized fundamental matrix* $\phi(t)$ equal to the flow map, with $\phi(t_0) = I$.

- In the autonomous case $\dot{x} = Ax$ solutions can be practically constructed.

(i) **Explicit Solution**

$$\phi(t) = e^{At} := \sum_{j=0}^{\infty} \frac{1}{j!}(At)^j \tag{3.22}$$

With $0! = 1$ and $0^0 = I$. We can verify that this is indeed a solution

$$\dot{\phi}(t) = \sum_{j=1}^{\infty} \frac{1}{(j-1)!}(At)^{j-1}A = A\sum_{j=0}^{\infty} \frac{1}{j!}(At)^j = Ae^{At} = A\phi(t). \tag{3.23}$$

Where we used that $A$ commutes with its powers in the second equality. We now have that each column of $\phi(t)$ satisfies $\dot{y} = Ay$.

*Remark* 3.6. For a scalar ODE $\dot{y} = a(t)y$ for $y \in \mathbb{R}$, the solution is known $y(t) = e^{\int_{t_0}^{t} a(s)ds} y_0$. However, this does not extend to the higher dimensional $\dot{y} = A(t)y$. In fact, in general, $\phi(t) = e^{\int_{t_0}^{t} A(s)ds}$ is <u>not</u> a solution. We can check this

$$\dot{\phi} = \sum_{j=0}^{\infty} \frac{1}{j!}\frac{d}{dt}\left(\int_{t_0}^{t} A(s)ds\right)^j = \sum_{j=1}^{\infty} \frac{1}{(j-1)!}\left(\int_{t_0}^{t} A(s)ds\right)^{j-1} A(t) \neq A(t)\phi(t). \tag{3.24}$$

The nonequality holds as $A(t)$ does not generally commute with $\int A(s)ds$.

(ii) **Solution from eigenfunctions** If we have an autonomous system, we can solve the ODE without an infinite series. We have

$$\dot{y} = Ay, \quad y \in \mathbb{R}^n, \quad y(0) = y_0. \tag{3.25}$$

Substituting $\varphi(t) = e^{\lambda t}s$ for $\lambda \in \mathbb{C}$ and $s \in \mathbb{C}^n$ into (3.25) yields

$$\lambda s = As \implies (A - \lambda I)s = 0 \iff \det(A - \lambda I) = 0. \tag{3.26}$$

Therefore $\lambda$ must be an eigenvalue of $A$ and $s$ must be the corresponding eigenvector. We call $\det(A-\lambda I)$ the *characteristic equation* of $A$. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues and $s_1, \ldots, s_n$ be the corresponding eigenvectors. In the case that some eigenvalues are repeated, some of the $s_i$ may be generalized eigenvectors. We then have two cases.

(a) $A$ is semisimple, i.e. the eigenvectors are linearly independent (which is always the case if the $\lambda_i$ all have algebraic multiplicity of one). Then we have the solution

$$y(t) = \sum_{i=1}^{n} c_i e^{\lambda_i t} s_i = \sum_{j=1}^{n} c_j e^{(\mathrm{Re}\lambda_j)t} e^{i(\mathrm{Im}\lambda_j)t} s_j. \tag{3.27}$$

Where we used $\lambda_j = \mathrm{Re}\lambda_j + i\mathrm{Im}\lambda_j$.

(b) $A$ is not semisimple, i.e. has repeated eigenvalues (but not enough linearly independent eigenvectors). Then we assume that $\lambda_k$ has algebraic multiplicity $a_k > g_k$, where $a_k$ measures the multiplicity of $\lambda_k$ as a root of $\det(A - \lambda I) = 0$, and $g_k$ is the number of linearly independent eigenvectors for $\lambda_k$, also called the *geometric muliplicity* of $\lambda_k$. Even in this case, $\lambda_k$ gives rise to $a_k$ linearly independent solutions of the form

$$\underbrace{P_0}_{=s_k} e^{\lambda_k t},\ P_1(t)e^{\lambda_k t},\ P_2(t)e^{\lambda_k t}, \ldots, P_{a_{k-1}}(t)e^{\lambda_k t} \tag{3.28}$$

where $P_j(t)$ is a vector polynomial of $t$ of order $j$ or less.

## 3.4 Stability of fixed points in autonomous linear systems

First we note that we can bound our solution

$$\|y(t)\| = \|\phi(t)y_0\| \leq \underbrace{\|\phi(t)\|}_{\text{Operator norm}} \|y_0\| \leq Ce^{\mu t}\|y_0\|. \tag{3.29}$$

Where $\mu = \max_j(\mathrm{Re}\lambda_j) + \nu$, with $\nu > 0$, as small as needed, provided we increase $C$ appropriately. If $A$ is semisimple, then $\nu = 0$ can be selected.

**Theorem 3.7** (Stability of fixed points in linear systems)**.** *Given $y = 0$ a fixed point of the linear system $\dot{y} = Ay$ with $A \in \mathbb{R}^{n\times n}$ the following statements hold:*

*(i) Assume that $Re\lambda_j < 0$ for all $j$. Then $y = 0$ is asymptotically stable.*

*(ii) Assume that $Re\lambda_j \leq 0$ for all $j$, and for all $\lambda_k$ with $Re\lambda_k = 0$ we have $a_k = g_k$. Then $y = 0$ is stable.*

*(iii) Assume there exists a $k$ such that $Re\lambda_k > 0$. Then $y = 0$ is unstable.*

*These scenarios are illustrated in Fig. 3.8.*

*Proof.* (i) Pick $\varepsilon > 0$, and select $\nu > 0$ small, such that $\mu < 0$. Then pick $C > 0$ such that (3.29) holds, and let $\delta = \frac{\varepsilon}{C}$. This implies (since $\|y_0\| \leq \delta$) that

$$\|y(t)\| \leq \varepsilon e^{\mu t} \leq \varepsilon, \tag{3.30}$$

and

$$\|y(t)\| \leq \varepsilon e^{\mu t} \xrightarrow{t\to\infty} 0. \tag{3.31}$$

Where the limit holds as $\mu < 0$.

(ii) Again choose $\delta = \frac{\varepsilon}{C}$ and note that $\mu = \max_j(\mathrm{Re}\lambda_j) + \nu = 0 + \nu = 0$ ($\nu = 0$ as $a_k = g_k$). Then stability follows by (3.29). However, asymptotic stability does not hold, as $\varphi(t) = Ce^{i(\mathrm{Im}\lambda_j)t}$ solutions exist.

(iii) There exists a solution of the form

$$\varphi(t) = C_k e^{\lambda_k t} s_k = C_k e^{(\mathrm{Re}\lambda_k)t} e^{i(\mathrm{Im}\lambda_k)t} s_k. \tag{3.32}$$

In turn this implies

$$\|\varphi(t)\| = C_k e^{(\mathrm{Re}\lambda_k)t} \|s_k\| \xrightarrow{t\to\infty} \infty. \tag{3.33}$$

□

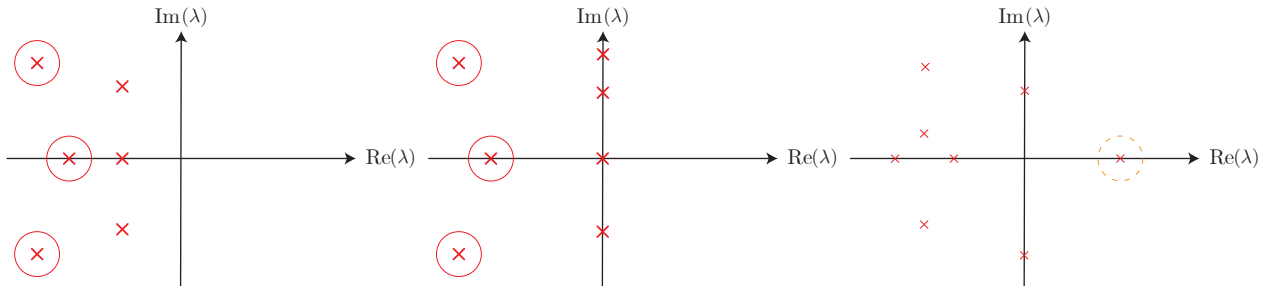


Figure 3.8: Eigenvalue arrangements for scenarios (i), (ii), and (iii) (from left to right) in Theorem 3.7

*Example* 3.5 (Stability analysis of 2 degrees of freedom coupled oscillators). Given a rectangular mass $m$ with a spring of stiffness coefficient $k$ attached to each side extending to fixed walls in each cardinal direction. We want to know the stability of the equilibrium where all of the springs are equally extended. This dynamical system is depicted in Fig. 3.9. First, note that this is a conservative system, i.e. $E = \text{const}$. Next we transform the coordinates so that the equations of motion can be brought into the form of an ODE for this dynamical system

$$x = \begin{pmatrix} x \\ \dot{x} \\ y \\ \dot{y} \end{pmatrix}. \tag{3.34}$$

Thus we have a 4-dimensional, nonlinear, system of ODEs. We now linearize this at the fixed point $(x, y) = (0, 0)$, i.e. $\dot{x} = Ax$ with $x \in \mathbb{R}^n$.

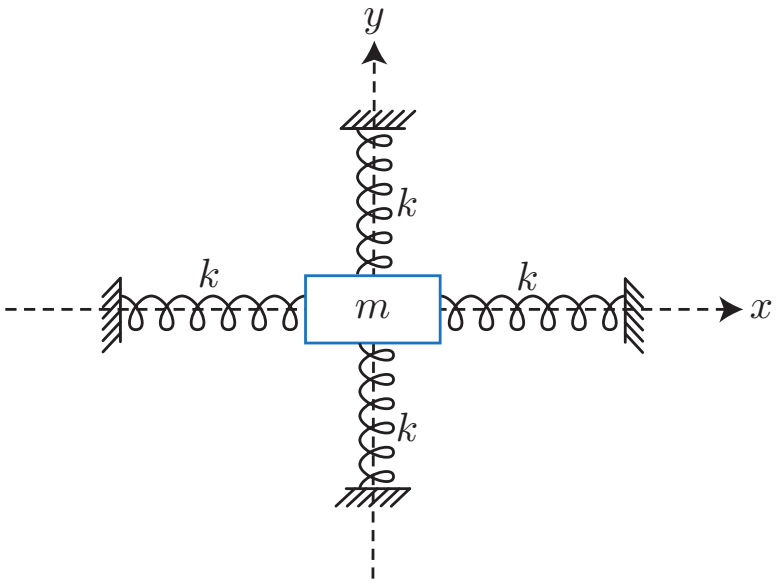


Figure 3.9: Arrangement of coupled oscillators with rectangular mass in the middle.

The system exhibits full spatial symmetry in $x$ and $y$, hence the eigenmodes will be the same in the $x$ and $y$ directions. This means we have repeated pairs of purely imaginary eigenvalues for $A$: $\lambda_{1,2} = \lambda_{3,4} = \pm i\omega$. It is clear that scenarios (i) and (iii) of Theorem 3.7 do not apply to the linearized ODE. So we need to check if (ii) applies.

We have that $\mathrm{Re}\lambda_k = 0$ for $k = 1, 2, 3, 4$. Also $a_k = 2$ for $k = 1, 2, 3, 4$. Now assume $g_k < 2$. Then there would exists solutions of the form $te^{\pm i\omega t}s_k$, but this would contradict the conservation of energy, as either the (nonnegative) kinetic energy and/or the (nonnegative) potential energy would grow unbounded. Hence, the total energy could not be conserved. Therefore we know that $g_k = a_k$ and we can apply (ii) to find $x = y = 0$ is Lyapunov stable for the linearized system. What does this imply for the nonlinear system?

**Exercise 3.1.** *Consider a discrete dynamical system given by the iterated mapping*

$$x_{k+1} = f(x_k); \quad f : \mathbb{R}^n \to \mathbb{R}^n, \ x_k \in \mathbb{R}^n. \tag{3.35}$$

*Assume that $x = p$ is a fixed point for the mapping, i.e., $p = f(p)$.*

*(a) Derive a linearized mapping of the form*

$$y_{k+1} = Ay_k \tag{3.36}$$

*to describe the discrete dynamics in the vicinity of the fixed point.*

*(b) Assume that $A$ has eigenvalues $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ with corresponding $n$ linearly independent eigenvectors $s_1, \dots, s_n \in \mathbb{C}^n$. Show that the general solution of (3.36) is of the form*

$$y_k = c_1\varphi_1(k) + \cdots + c_n\varphi_n(k), \qquad \varphi_i(k) = \lambda_i^k s_i. \tag{3.37}$$

*(c) Formulate a definition of stability, asymptotic stability, and instability for the $y = 0$ fixed point of (3.36).*

*(d) Using (3.37), find a sufficient and necessary condition for the asymptotic stability you have defined in (c).*

**Exercise 3.2.** *Consider a ball of mass $m$ that slides on a rotating hoop (see Fig. 3.10).*

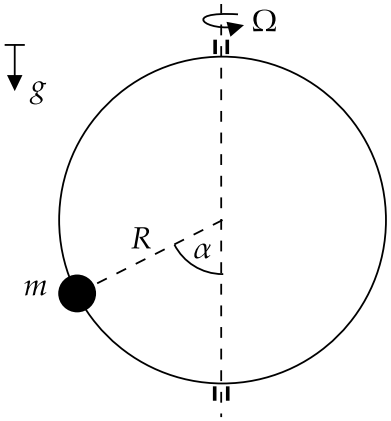


Figure 3.10: Mass on a hoop

*The angular velocity of the hoop is $\Omega$, the viscous friction coefficient between the hoop and the ball is $b$, and the constant of gravity in $g$. The equation of motion for the sliding ball is given by*

$$mR^2\ddot{\alpha} + bR^2\dot{\alpha} + mR^2(g/R - \Omega^2\cos(\alpha))\sin(\alpha) = 0. \tag{3.38}$$

*(a) Plot the location of equilibria of the ball as a function of the non-dimensionalized rotation parameter $\nu = R\Omega^2/g$.*

*(b) Using linearization, determine the stability type of the different equilibrium branches on the plot. Identify the critical angular velocity at which a bifurcation of equilibria occurs.*

## 3.5 Stability of fixed points in nonlinear systems

Following the previous example, we would like to know what information about the stability of fixed points of nonlinear systems we can derive from the linearized system. The full nonlinear system is

$$\dot{x} = f(x), \quad f(x_0) = 0, \quad x \in \mathbb{R}^n, \quad f \in \mathcal{C}^1. \tag{3.39}$$

And its linearization at the fixed point $x_0$

$$\dot{y} = Df(x_0)y, \quad y \in \mathbb{R}^n, \quad Df(x_0) \in \mathbb{R}^{n\times n}. \tag{3.40}$$

We would like to conclude that the linearized dynamics are qualitatively similar to the nonlinear dynamics. In order to study if this is the case, we have to formalize *similar* mathematically.

**Definition 3.7** ($\mathcal{C}^k$ equivalence of dynamical systems)**.** Consider two autonomous dynamical systems:

(i)

$$\dot{x} = f(x), \quad x \in \mathbb{R}^n, \quad f \in \mathcal{C}^1; \quad F^t : x_0 \mapsto x(t; x_0). \tag{3.41}$$

(ii)

$$\dot{x} = g(x), \quad x \in \mathbb{R}^n, \quad g \in \mathcal{C}^1; \quad G^t : x_0 \mapsto x(t; x_0). \tag{3.42}$$

The two dynamical systems are $\mathcal{C}^k$ *equivalent*, for $k \in \mathbb{N}$, on an open set $U \subset \mathbb{R}^n$, if there exists a $\mathcal{C}^k$ diffeomorphism $h : U \to U$ that maps orbits of (i) into orbits of (ii), while preserving the orientation but not necessarily the exact parameterization of the orbit by time. Specifically for all $x \in U$, any $t_1 \in \mathbb{R}$ there exists a $t_2 \in \mathbb{R}$ such that

$$\boxed{h(F^{t_1}(x)) = G^{t_2}(h(x)).} \tag{3.43}$$

$h : U \to U$ does this for all $x \in U$ in a $\mathcal{C}^k$ fashion. This equivalence through a function $h$ is demonstrated in Fig 3.11.

**Proposition 3.8** (Restrictiveness of smooth equivalence)**.** *Consider two $\mathcal{C}^k$-equivalent dynamical systems for $k > 0$ with the right hand sides $f$ and $g$, as in the definitions above. Let $x_0$ be a fixed point of $F$ and $y_0$ a fixed point of $G$ such that these correspond to each other under the $\mathcal{C}^k$ equivalence, i.e. $y_0 = h(x_0)$. The eigenvalues of $Df(x_0)$ must be constant, positive multiples of the eigenvalues of $Dg(y_0) = Dg(h(x_0))$.*

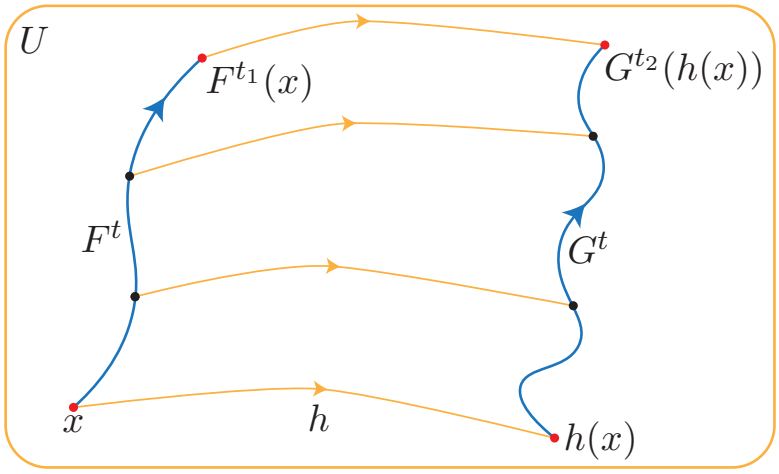


Figure 3.11: The function $h$ mapping the orbits of the dynamical system describing $F$ into the system describing $G$.

*Proof.* By the $\mathcal{C}^k$ equivalence we have that there exists a $\tau(x,t)$ with $\frac{d\tau}{dt} > 0$ such that $h(F^t(x)) = G^{\tau(x,t)}(h(x))$. Further assume that $\tau$ is at least continuously differentiable in its arguments. Differentiating the previous equality with respect to $x$ yields

$$DhDF^t = \dot{G}^\tau \frac{d\tau}{dx} + DG^\tau Dh. \tag{3.44}$$

At a fixed point $x_0$ we have that $\dot{G}^\tau(x_0) = g(G^\tau(x_0)) = 0$. This implies that the first term on the right hand side of the differentiated equality is equal to 0. Now denote the linearizations of $f$ and $g$ at the fixed point by $A = Df(x_0)$ and $B = Dg(h(x_0))$ respectively. Then the matrices $e^{At}$ and $e^{B\tau}$ must be similar matrices, in turn implying that $At$ and $B\tau(x_0,t)$ must have the same spectrum, i.e.

$$\frac{\lambda_i(B)}{\lambda_i(A)} = \frac{t}{\tau(x_0,t)} = C = \text{const.} > 0. \tag{3.45}$$

Therefore the eigenvalues of the linearizations are constant, positive multiples of each other. □

**Definition 3.8** ($\mathcal{C}^k$ conjugacy)**.** Consider the same two dynamical systems as before. The two dynamical systems are $\mathcal{C}^k$ *conjugate* if there exists a $\mathcal{C}^k$ diffeomorphism $h : U \to U$ that maps orbits of (i) into (ii), while preserving orientation and the parameterization of the orbit by time. Specifically for all $x \in U$ and $t \in \mathbb{R}$ we have

$$\boxed{h(F^t(x)) = G^t(h(x)).} \tag{3.46}$$

**Proposition 3.9** (Restrictiveness of smooth conjugacy)**.** *Consider two $\mathcal{C}^k$ conjugate systems, as in the definition above. In addition to the previous proposition, now the linearizations of the two systems will have the same spectra at each fixed point.*

*Proof.* The $\mathcal{C}^k$ conjugacy requires preserving the parameterizations of the orbits, i.e. $\tau(x,t) = t$. Hence, this is just a specification of the previous proof. □

**Definition 3.9** (Topological equivalence)**.** For $k = 0$, $\mathcal{C}^k$ equivalence is also called *topological equivalence*. In this case, a continuous, invertible deformation takes orbits of one system into the orbits of the other. Under these conditions, $h : U \to U$ is called a *homeomorhpism*.

*Example* 3.6 (Topologically equivalent linear systems for $n = 2$). To illustrate the meaning of topological equivalence, Fig. 3.12 shows three linear systems ($\dot{x} = Ax$ for $x \in \mathbb{R}^2$) which are topologically equivalent.

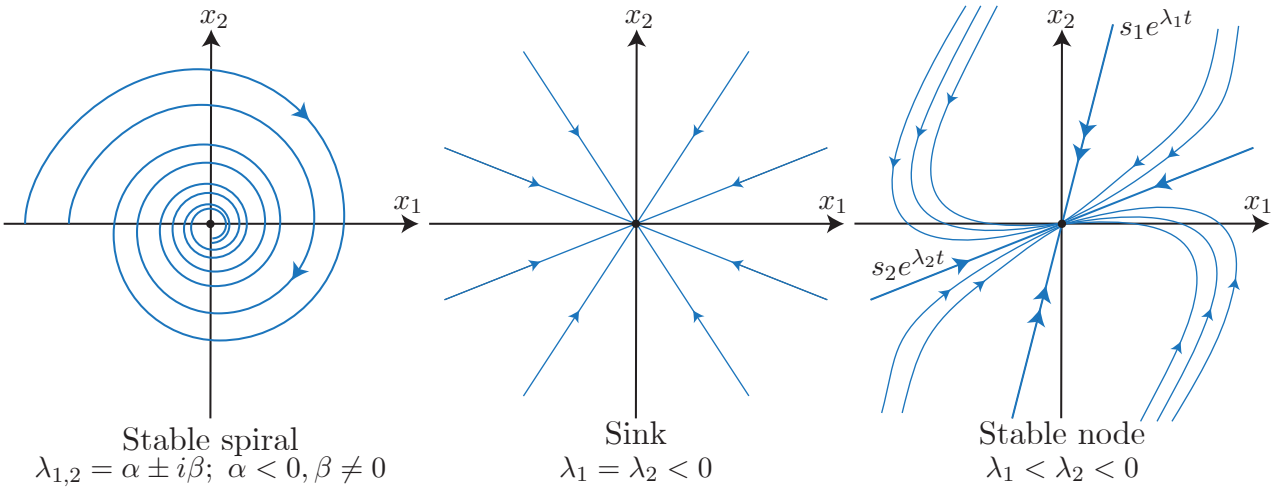


Figure 3.12: Three topologically equivalent 2-dimensional linear systems. Left: The stable spiral. Middle: The sink. Left: The stable node.

The stable spiral has the eigenvalues $\lambda_{1,2} = \alpha \pm i\beta$ for $\alpha < 0$ and $\beta \neq 0$. The sink has the eigenvalues $\lambda_1 = \lambda_2 < 0$. and The stable node has the eigenvalues $\lambda_1 < \lambda_2 < 0$. Note here that the number of eigenvalues $\lambda_i$ with $\mathrm{Re}\lambda_i < 0$, $\mathrm{Re}\lambda_i = 0$, and $\mathrm{Re}\lambda_i > 0$ is the same in all three of these cases, namely for each system the real part of both eigenvalues are less than 0.

*Remark* 3.10. The above systems are not $\mathcal{C}^k$ equivalent or conjugate, as the spectra differ by more than a multiplicative constant.

*Example* 3.7 (Topologically inequivalent linear systems for $n = 2$). As a counter example, we now present three linear systems which are not topologically equivalent. The stable spiral (from

before), the unstable spiral, and the saddle. The unstable spiral has the eigenvalues $\lambda_{1,2} = \alpha \pm i\beta$ for $\alpha > 0$ and $\beta \neq 0$ (note the different sign for $\alpha$). The saddle has the eigenvalues $\lambda_1 < 0 < \lambda_2$. These systems are depicted in Fig 3.13.

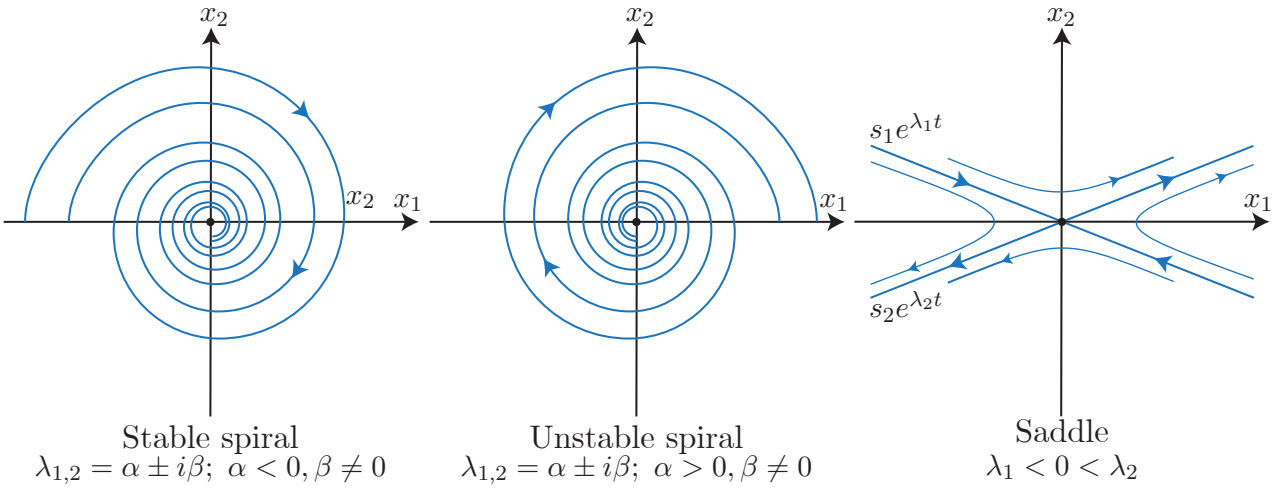


Figure 3.13: Three 2-dimensional linear systems which are not topologically equivalent. Left: The stable spiral. Middle: The unstable spiral. Right: The saddle.

Note here that the eigenvalue configurations in terms of the number of $\lambda_i$ with real part less than 0 are different in each case. Building on the role of the eigenvalue configuration we noted in the previous examples, we introduce the concept of a hyperbolic fixed point.

**Definition 3.10** (Hyperbolic fixed point)**.** We call the fixed point $x = x_0$ a *hyperbolic fixed point* of (3.39) if each of the eigenvalues $\lambda_i$ of its linearization (3.40) satisfy

$$\boxed{\mathrm{Re}\lambda_i \neq 0.} \tag{3.47}$$

Geometrically the eigenvalue configuration of a hyperbolic fixed point is shown in Fig. 3.14.

**Proposition 3.11.** *Under small perturbations to the nonlinear system, the linearized stability type of the hyperbolic fixed point is preserved.*

Before proving this result, recall the Implicit Function Theorem (without proof).

**Theorem 3.12** (Implicit Function Theorem ($n+1$ dimensional case))**.** *For a function $F : \mathbb{R}^{n+1} \to \mathbb{R}^n$ which is $\mathcal{C}^1$, if $F(x_0, y_0) = 0$ and the Jacobian $D_x F(x_0, y_0)$ is nonsingular (invertible), then there exists a nearby solution to $F(x, y) = 0$, for $x_y = x_0 + \mathcal{O}(|y - y_0|)$. Further $x_y$ is as smooth in $y$ as $F(x, y)$.*

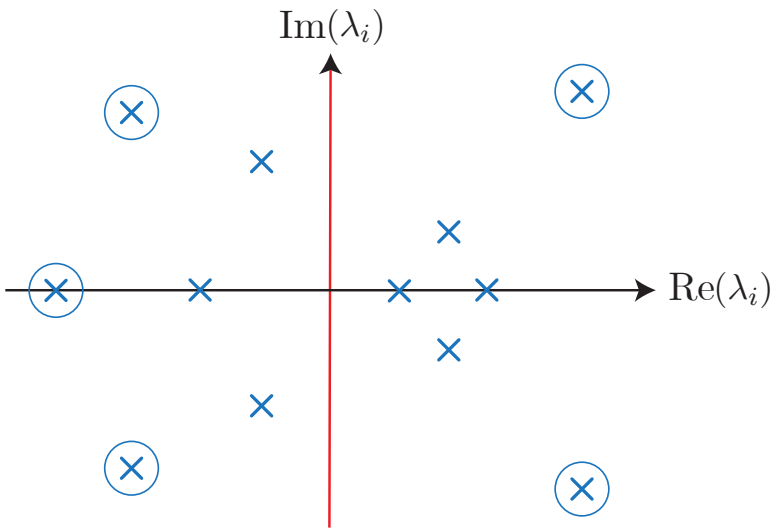


Figure 3.14: The eigenvalue configuration of a hyperbolic fixed point, i.e. no eigenvalues are on the imaginary axis (red).

*Proof (Proposition).* Add a small perturbation to (3.41) i.e.

$$\dot{x} = f(x) + \varepsilon g(x); \quad |\varepsilon| \ll 1, \quad f(x_0) = 0. \tag{3.48}$$

Now we ask if the perturbed system has a fixed point $x_\varepsilon$ near $x_0$. We frame this in terms of the implicit function theorem

$$F(x, \varepsilon) = f(x) + \varepsilon g(x) \stackrel{?}{=} 0; \quad F(x_0, 0) = 0; \quad x \in \mathbb{R}^n, \quad F : \mathbb{R}^{n+1} \to \mathbb{R}^1. \tag{3.49}$$

We check that $D_x F(x_0, 0)$ is nonsingular exactly when $Df(x_0)$ is; this is fulfilled as we have no zero eigenvalues. The linearization at the perturbed fixed point takes the form

$$\dot{y} = D\left[f(x) + \varepsilon g(x)\right]\big|_{x=x_\varepsilon} y = \left[Df(x_0 + \mathcal{O}(\varepsilon)) + \varepsilon Dg(x_0 + \mathcal{O}(\varepsilon))\right] y = \underbrace{\left[Df(x_0) + \mathcal{O}(\varepsilon)\right]}_{=A_\varepsilon} y. \tag{3.50}$$

In the last equality we used the Taylor expansion in $\varepsilon$. We have that the roots of $\det(A_\varepsilon - \lambda I) = 0$ depend continuously on the parameter $\varepsilon$. These roots correspond to the eigenvalues of $A_\varepsilon$. Therefore, the roots stay within an $\mathcal{O}(\varepsilon)$ neighborhood of the eigenvalues of $Df(x_0)$ (see Fig. 3.15). Hence we have that the eigenvalue configuration is unchanged for small enough $\varepsilon$. □

*Remark* 3.13. In the above proof, not only does the hyperbolicity of fixed points remain preserved, but also the stability type.

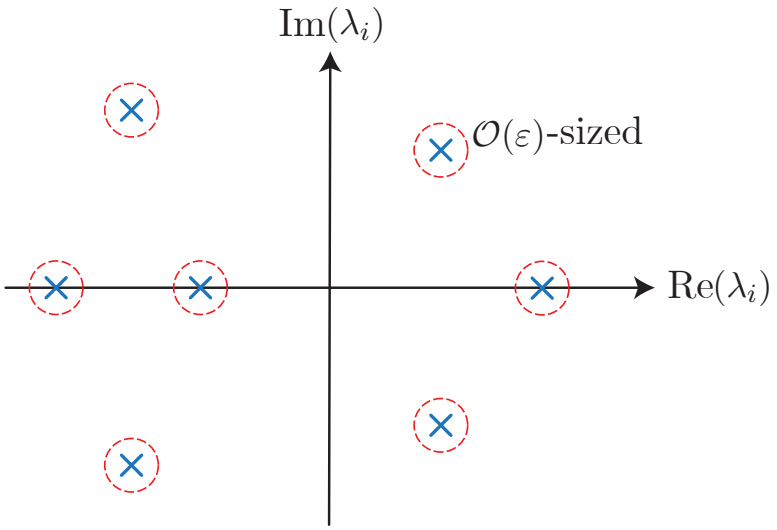


Figure 3.15: The eigenvalue configuration the $\mathcal{O}(\varepsilon)$ neighborhood (red) drawn around each eigenvalue (blue).

Meanwhile, for nonhyperbolic fixed points, this is not the case, and the smallest perturbation may change their stability type. This is due to the fact, that no matter how small the scale of the perturbation ($\varepsilon$) the $\mathcal{O}(\varepsilon)$ ball around eigenvalues on the imaginary axis will always intersect with $\mathbb{C} - \{\text{Im}\lambda_i = 0\}$ (i.e. points which are not on the imaginary axis).

Now we would like to connect the preservation of stability type under nonlinear perturbation to analyzing the stability type of fixed points of nonlinear dynamical systems based on their linearization.

**Theorem 3.14** (Hartman-Grobman)**.** *If the fixed point $x_0$ of the nonlinear system (3.39) is hyperbolic, then the linearization (3.40) is topologically equivalent to the nonlinear system in a neighborhood of $x_0$.*

***Consequence:*** *For hyperbolic fixed points, linearization predicts the correct stability type and orbit geometry near $x_0$.*

Now we would like to apply this to the pendulum to systematically derive the stability type of its fixed points.

*Example* 3.8 (Stability analysis of the pendulum via Hartman-Grobman). Recall the transformed ODE for the pendulum

$$\dot{x} = f(x) = \begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = -\sin(x_1) \end{cases}; \quad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}. \tag{3.51}$$

We have two fixed points $p = (\pi, 0)$ and $q = (0, 0)$. First we analyze the stability of the fixed point $p$. The differential of $f$ at a point $a$ is

$$Df(a) = \begin{pmatrix} 0 & 1 \\ -\cos(a_1) & 0 \end{pmatrix}. \tag{3.52}$$

Start by linearizing at $p$

$$\dot{y} = Ay; \quad A = Df(p) = \begin{pmatrix} 0 & 1 \\ -\cos(x_1) & 0 \end{pmatrix}_{x=p} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{3.53}$$

Now we have to check if $p$ is hyperbolic

$$\det(A - \lambda I) = \lambda^2 - 1 = 0 \implies \lambda_{1,2} = \pm 1; \; s_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \; s_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \tag{3.54}$$

Neither of the eigenvalues lie on the imaginary axis, so $p$ is hyperbolic. This allows us to move between the nonlinear and linearized system for the stability analysis without compromising our results. We find the linearized dynamics to be

$$y(t) = C_1 e^{-t} s_1 + C_2 e^{t} s_2. \tag{3.55}$$

Using the initial conditions $y_0$ the trajectory can be expressed as

$$y(t) = F^t y_0, \tag{3.56}$$

where $F^t$ is the normalized fundamental matrix solution. We can now fully describe the phase portrait of the linearization. The *stable subspace* $E^S$ is $\text{span}\{s_1\} = \{y_0 : \; F^t y_0 \xrightarrow{t\to\infty} 0\}$. The unstable subspace $E^U$ is $\text{span}\{s_2\} = \{y_0 : \; F^t y_0 \xrightarrow{t\to-\infty} 0\}$. The phase portrait near the fixed point $p$ is illustrated in Fig. 3.16

The nonlinear phase portrait is topologically equivalent to the linear one. Further we can define the stable and unstable manifolds of $p$ for the nonlinear system. We designate $F^t(\cdot)$ to be the flow map for the nonlinear system after this point. The *stable manifold* of $p$ is

$$\boxed{W^S = \{x_0 : \; F^t(x_0) \xrightarrow{t\to\infty} p\}.} \tag{3.57}$$

and the *unstable manifold* of $p$

$$\boxed{W^U = \{x_0 : \; F^t(x_0) \xrightarrow{t\to-\infty} p\}.} \tag{3.58}$$

Both of these are $\mathcal{C}^0$ curves through $p$ and their existence follows from the Hartman-Grobman theorem. These manifolds are shown in the nonlinear phase portrait around $p$ in Fig. 3.17.

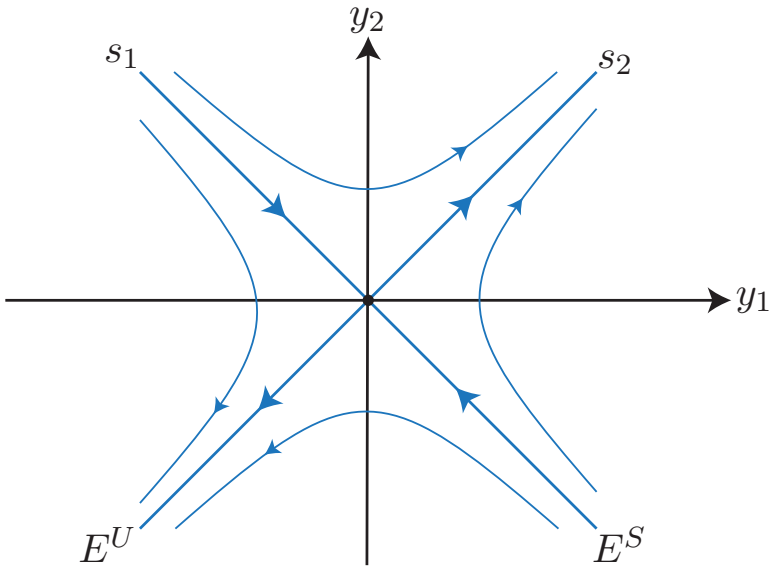


Figure 3.16: The phase portrait of the linearized pendulum in a neighborhood around $p$.

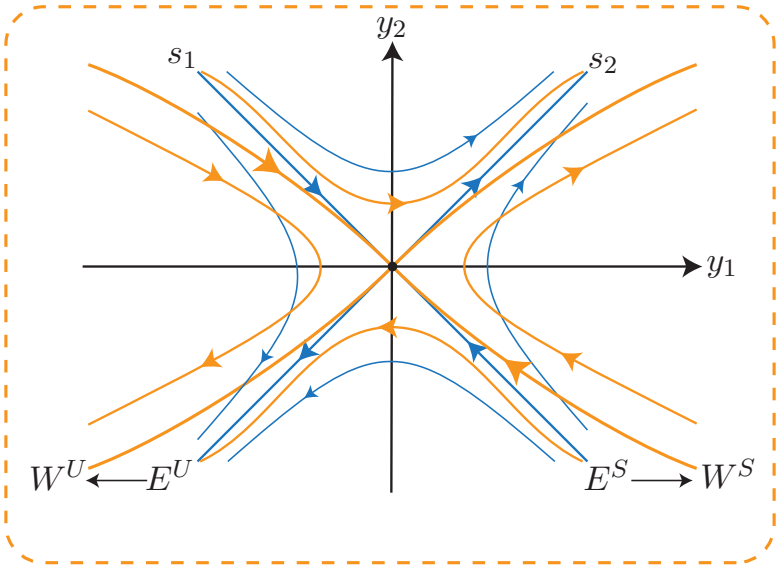


Figure 3.17: The phase portrait of the pendulum on a neighborhood around $p$ with the stable and unstable manifolds of the nonlinear system as well as the stable and unstable spaces of the linearization.

Next we analyze the stability of the fixed point $q$. Once again, our first step is to linearize

$$\dot{y} = Ay, \quad A = Df(q) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}; \quad \det(A - \lambda I) = 0. \tag{3.59}$$

From here, we see that the determinant is equal $\lambda^2 + 1 = 0$, yielding the roots $\lambda_{1,2} = \pm i$, i.e. the fixed point is not hyperbolic. Thus the linearized dynamics is inconclusive for the nonlinear system. In this particular case, $q$ turns out to be stable by the definition of Lyapunov stability. Later we will use another approach to show this directly.

In the last example, the importance of hyperbolicity was not accentuated, as the latter fixed point had the same stability type in the linearized system as in the full system. This leads us to question if there are cases where the stability type between the linear and nonlinear systems is not preserved.

*Example* 3.9 (Criticality of hyperbolicity in Hartman-Grobman). Let the dynamical system be

$$\dot{x} = ax^3, \quad x \in \mathbb{R}, \quad a \neq 0. \tag{3.60}$$

This system has a fixed point at $x = 0$, linearizing here gives

$$A = 3ax^2\big|_{x=0} = 0 \implies \dot{y} = 0y = 0. \tag{3.61}$$

We have a single root $\lambda_1 = 0$, hence $x = 0$ is a nonhyperbolic fixed point. Disregarding this fact, we may be inclined to conclude that $x = 0$ is a stable fixed point, since $y = 0$ is trivially a fixed point of the linearization $\dot{y} = 0$. This is not the case, as we can see by analyzing the full nonlinear dynamics for $a > 0$ as in Fig. 3.18, where we observe that $x = 0$ is an unstable fixed point.

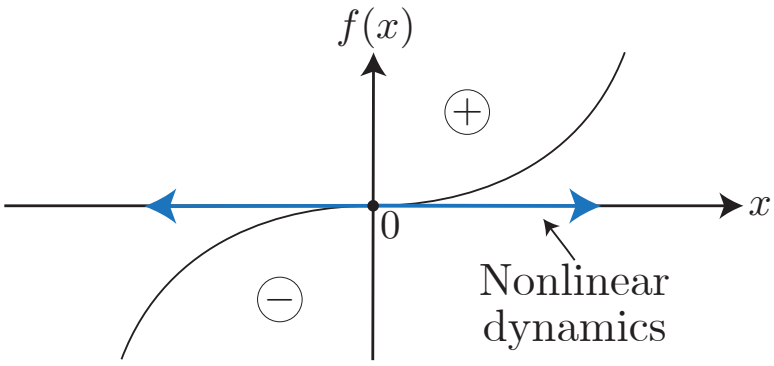


Figure 3.18: Nonlinear dynamics for the dynamical system $\dot{x} = ax^3$ with $a > 0$.

Now that we have understood how to use the Hartman-Grobman theorem, we would like to to be able to definitely conclude the stability type of a fixed point, once we have the linearization. To achieve this, we require a sufficient and necessary criterion for all of the eigenvalues of the linearized system to be left of the imaginary axis, i.e. $\mathrm{Re}\lambda_i < 0$ for all $i$.

**Theorem 3.15** (Routh-Hurwitz)**.** *Consider the polynomial*

$$a_n\lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0 = 0. \tag{3.62}$$

*Without loss of generality assume $a_0 > 0$, if $a_0 < 0$ then multiply by $-1$ and if $a_0 = 0$ then $\lambda = 0$ is a root and therefore we cannot have asymptotic stability. Next, define the following series of subdeterminants*

$$D_0 = a_0,\ D_1 = a_1,\ D_2 = \begin{vmatrix} a_1 & a_0 \\ a_3 & a_2 \end{vmatrix},\ D_3 = \begin{vmatrix} a_1 & a_0 & 0 \\ a_3 & a_2 & a_1 \\ a_5 & a_4 & a_3 \end{vmatrix}, \ldots,\ D_n = \begin{vmatrix} a_1 & a_0 & 0 & \ldots & 0 \\ a_3 & a_2 & a_1 & \ldots & 0 \\ \vdots & & & & \vdots \\ 0 & \ldots & a_n & a_{n-1} & a_{n-2} \\ 0 & \ldots & & 0 & a_n \end{vmatrix} \tag{3.63}$$

*Then we have that if and only if for all $i$ $D_i > 0$ then $Re\lambda_i < 0$ for all $i$.*

*A weaker neccesary condition is that if for all $i$ $Re\lambda_i < 0$ then $a_i > 0$ for all $i$. Therefore if there exists an $a_k < 0$, we know immediately that the fixed point cannot be asymptotically stable as not all of the $Re\lambda_i$ can be strictly negative.*

*Remark* 3.16. For a given $i$ we construct the matrix used for calculating $D_i$ as follows: write the elements $a_1, \ldots, a_i$ along the diagonal, then in each row $k$ write the $a_j$ in descending index order such that $a_k$ aligns with the placement inherited from us writing along the diagonal. The leftover spaces are filled with zeros.

*Remark* 3.17. Adolf Hurwitz discovered this criterion independently of Edward Routh in 1895 while holding a chair at the ETH.

*Example* 3.10 (Applying the Routh-Hurwitz criterion). Given the polynomial

$$a_3\lambda^3 + a_2\lambda^2 + a_1\lambda + a_0 = 0. \tag{3.64}$$

The Routh-Hurwitz criterion is

$$D_0 = a_0 > 0;\ D_1 = a_1 > 0;\ D_2 = a_1a_2 - a_0a_3 > 0;\ D_3 = a_3D_2 > 0. \tag{3.65}$$

Therefore

$$\boxed{a_0 > 0,\ a_1 > 0,\ a_1a_2 - a_0a_3 > 0,\ a_3 > 0} \tag{3.66}$$

forms a sufficient and necessary condition for asymptotic stability ($a_i > 0$ follows from here) for $n = 3$.

**Exercise 3.3.** *The first three modes of a convecing fluid motion in a two-dimensional layer heated from below are described by the famous* Lorenz equations

$$\dot{x} = a(y - x), \tag{3.67}$$

$$\dot{y} = bx - y - xz, \tag{3.68}$$

$$\dot{z} = xy - cz. \tag{3.69}$$

*Here $a > 0$ denotes the Prandtl number, $b > 0$ is the Rayleigh number, and $c > 0$ is the aspect ratio. Lorenz's original assumption is that $a > 1 + c$.*

*The above equations describe the evolution of the amplitudes of one velocity mode and two temperature modes. As a paradigm of chaotic dynamics, this system has inspired much of the development of the modern geometric theory of dynamical systems.*

*(a) Complicated dynamics arise when all possible equilibria of the system become unstable, and hence solutions cannot settle down to any simple steady state. Show that this is the case when*

$$b > \frac{a(3 + a + c)}{a - c - 1} \tag{3.70}$$

*Note: Note: A negative sign for one of the Routh-Hurwitz determinants actually implies instability, not just the lack of asymptotic stability.*

*(b) Solve the Lorenz equations numerically for $a = 10$, $b = 28$, and $c = 8/3$, choosing an initial condition close to $x = y = z = 0$. Plot the trajectory in three dimensions to show that it converges to a complicated surface, a* chaotic attractor*.*

A more refined linearization result comes from Sternberg's Theorem [Chicone, 1999].

**Theorem 3.18** (Sternberg)**.** *Assume the following*

*(i) The fixed point is asymptotically stable, i.e. $Re\lambda_j < 0$ for all $j = 1, \ldots, n$ (this also implies hyperbolicity);*

*(ii) $f \in \mathcal{C}^r$ for $r \in \mathbb{N} \cup \{\infty\}$ and $r > \frac{\max_j |Re\lambda_j|}{\min_j |Im\lambda_j|}$;*

*(iii) The system is nonresonant, i.e.*

$$\sum_{j=1}^{n} m_j \lambda_j \neq \lambda_k, \ \forall k \tag{3.71}$$

*for any sequence $m_j \in \mathbb{N}$ with $\sum_{j=1}^{n} m_j \geq 2$.*

*Then the following holds:* $\dot{x} = f(x)$ *with* $f(0) = 0$ *is locally* $\mathcal{C}^r$ *conjugate to its linear part* $\dot{y} = Ay$ *for* $A = Df(0)$.

*Example* 3.11 (Watt's centrifugal governor for steam engines). Now we put together everything built until now in the example of Watt's centrifugal governor for steam engines. Originally this system was used in mills in the 1700's, then it was adapted by Watt to the steam engine in 1788. This adaptation has been credited as a major factor in the industrial revolution, and is a first example of feedback control. The system is outlined in Fig. 3.19. The two masses (of mass

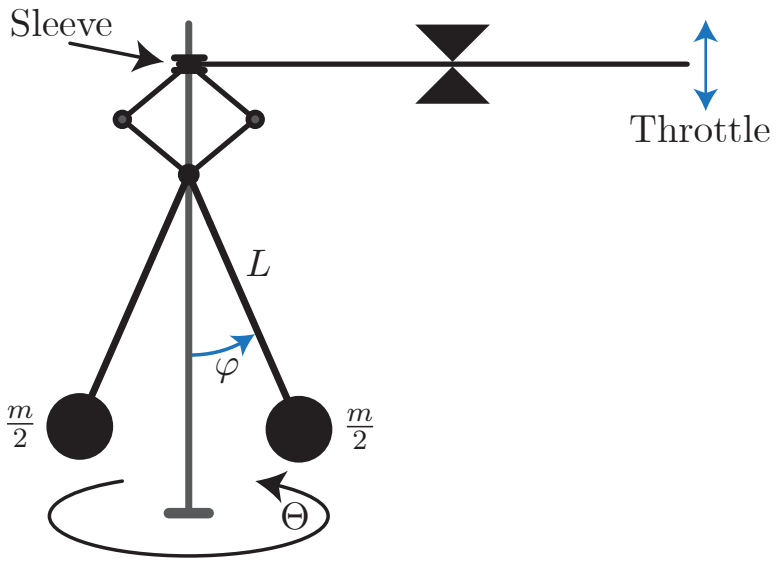


Figure 3.19: Schematic for Watt's centrifugal governor. The sleeve acts as a damper on the rotation about the spindle (drawn in grey).

$\frac{m}{2}$) rotate counter clockwise, and their position in radians is given by $\theta$, with their rotational velocity $\dot{\theta}$. The masses are attached by a rod of length $L$ and their deflection from the vertical position is measured by $\varphi$. Smaller $\dot{\theta}$ allowed for an increase in steam supply.

Following changes in the design, the systems suddenly became unstable. To address this Vyshnegradsky studied the root cause in 1877. We first derive the equation of motion. For the governor we use the equation of motion for a rotating hoop (with viscous damping coefficient $b$).

$$mL^2\ddot{\varphi} + +bL^2\dot{\varphi} + mL^2\left(\frac{g}{L} - \dot{\theta}^2\cos(\varphi)\right)\sin(\varphi) = 0; \quad b > 0. \tag{3.72}$$

Next, let $\omega$ denote the angular velocity of the steam engine, i.e. with the gear ratio $n$ we have $\dot{\theta} = n\omega$. Then we find

$$m\ddot{\varphi} = -b\dot{\varphi} - m\left(\frac{g}{L} - n^2w^2\cos(\varphi)\right)\sin(\varphi). \tag{3.73}$$

Now we derive the equation of motion for the steam engine. Denote the moment of inertia for the engine by $J$, the driving torque from the steam as $P_1$ and the constant load $P$, we obtain

$$J\dot{\omega} = P_1 - P. \tag{3.74}$$

In this case we have $P_1 = P^* + k\left(\cos(\varphi) - \cos(\varphi^*)\right)$ for the desired operation angle $\varphi^*$, the gain $k$, and $P*$ the value of $P$ at $\varphi^*$. Putting this together yields

$$J\dot{\omega} = k\cos(\varphi) - P_0; \quad P_0 = P - P^* + k\cos(\varphi^*). \tag{3.75}$$

Let $\dot{\varphi} = \Psi$ to transform into a three-dimensional set of equations (ODE)

$$\begin{cases} \dot{\varphi} = \Psi \\ \dot{\Psi} = -\dfrac{b}{m}\Psi - \left(\dfrac{g}{L} - n^2\omega^2\cos(\varphi)\right)\sin(\varphi)\,; \\ \dot{\omega} = \dfrac{k}{J}\cos(\varphi) - \dfrac{P_0}{J}. \end{cases} \quad x = \begin{pmatrix} \varphi \\ \Psi \\ \omega \end{pmatrix}. \tag{3.76}$$

Then our operation point is the fixed point $x_0$ of this system

$$f(x_0) = 0 \implies \Psi_0 = 0;\ \omega_0^2 = \frac{g}{Ln^2\cos(\varphi_0)};\ \cos(\varphi_0) = \frac{P_0}{k}. \tag{3.77}$$

If $\sin(\varphi_0) = 0$, we have an unphysical state and ignore this case. For simplification set $L = 1$, this could be formally achieved by nondimensionalizing the length $L$. Now we linearize at the fixed point $x_0$

$$\dot{y} = Ay; \quad A = Df(x_0) = \begin{pmatrix} 0 & 1 & 0 \\ n^2\omega^2\cos(2\varphi_0) - g\cos(\varphi_0) & -\frac{b}{m} & n^2\omega_0\sin(2\varphi_0) \\ -\frac{k}{J}\sin(\varphi_0) & 0 & 0 \end{pmatrix}. \tag{3.78}$$

With this we obtain the characteristic equation $\det(A - \lambda I) = 0$

$$\underbrace{1}_{a_3}\lambda^3 + \underbrace{\frac{b}{m}}_{a_2}\lambda^2 + \underbrace{g\frac{\sin^2(\varphi_0)}{\cos(\varphi_0)}}_{a_1}\lambda + \underbrace{g\frac{2k\sin^2(\varphi_0)}{J\omega_0}}_{a_0} = 0. \tag{3.79}$$

Now check the Routh-Hurwitz criterion for asymptotic stability:

(i) The necessary condition for $\mathrm{Re}\lambda_k < 0$ for all $k$: $a_j > 0$ for all $j$ is fulfilled.

(ii) Next we check the subdeterminants

$$D_0 = a_0 = g\frac{2k\sin^2(\varphi_0)}{J\omega_0} > 0; \tag{3.80a}$$

$$D_1 = a_1 = g\frac{\sin^2(\varphi_0)}{\cos(\varphi_0)} > 0; \tag{3.80b}$$

$$D_2 = \begin{vmatrix} a_1 & a_0 \\ a_3 & a_2 \end{vmatrix} = a_1a_2 - a_0a_3 = g\frac{b}{m}\frac{\sin^2(\varphi_0)}{\cos(\varphi_0)} - g\frac{2k\sin^2(\varphi_0)}{J\omega_0} > 0; \tag{3.80c}$$

$$D_3 > 0 \Longleftrightarrow D_2 > 0 \text{ and } a_3 = 1 > 0. \tag{3.80d}$$

The only actual condition for $x_0$ to be asymptotically stable is $D_2 > 0$

$$\frac{bJ}{m} > \frac{2P_0}{\omega_0}. \tag{3.81}$$

From the equation for the fixed points we know

$$P_0\omega_0^2 = \frac{gk}{Ln^2} = \text{const.} \tag{3.82a}$$

$$\omega_0^2 + 2P_0\omega_0\frac{d\omega_0}{dP_0} = 0. \tag{3.82b}$$

From the first equation, we realize that we must write $\omega_0 = \omega_0(P_0)$. To obtain the second equation, we derive the first equation (3.82a) with respect to $P_0$. Therefore we find

$$\frac{d\omega_0}{dP_0} = -\frac{\omega_0}{2P_0}. \tag{3.83}$$

Next, define the *non-uniformity of performance*

$$\nu = \left|\frac{d\omega_0}{dP_0}\right| = \frac{\omega_0}{2P_0}. \tag{3.84}$$

Then our criterion for asymptotic stability becomes

$$\frac{bJ}{m}\nu > 1. \tag{3.85}$$

To conclude, the following have harmful effects of stability

- Bigger engines which increase $m$
- Better machining of surfaces decreasing $b$
- Increased operating speed decreasing $J$
- Versatility in operation decreasing $\nu$.

These have the effect of pushing the eigenvalues to the right in the complex plane (towards the imaginary axis) as is illustrated in Fig 3.20.

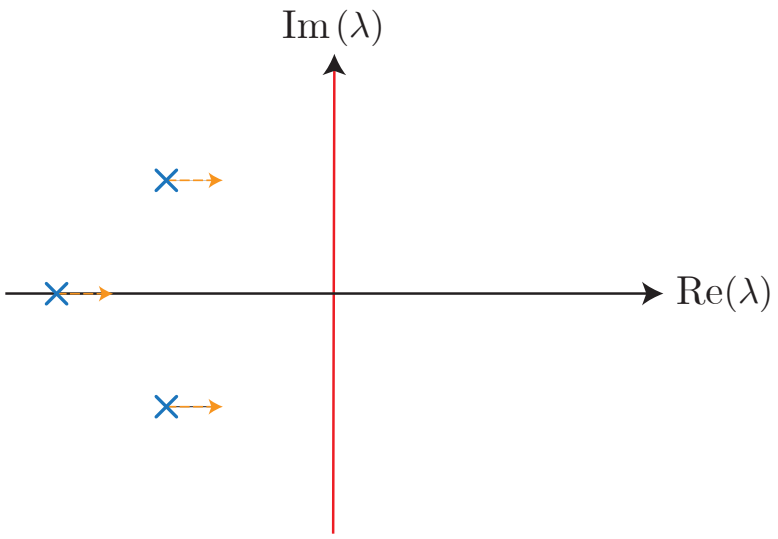


Figure 3.20: Change in the eigenvalue configuration as small modification are made to the Watt engine governor.

## 3.6 Data-driven linear modeling of dynamical systems

Given are $m+1$ vectors of $n$-dimensional observables $x_0, \ldots, x_m \in \mathbb{R}^n$ observed at $\Delta t$ intervals. An example of such data would be many different trajectories or snapshots of a single trajectory. Typically $m \gg 1$, i.e. there are a high number of observations.

Our objective is to find the best fitting linear dynamical system

$$\dot{x} = Ax; \quad x \in \mathbb{R}^n; \quad A \in \mathbb{R}^{n \times n}. \tag{3.86}$$

Since the data is given as discrete observations, this is equivalent to seeking the best $B = e^{A\Delta t}$ such that $x_{k+1} = Bx_k$ for $k = 1, \ldots, m-1$. To solve this we synthesize the data into matrices

$$X = \begin{bmatrix} x_0 \ldots x_{m-1} \end{bmatrix} \in \mathbb{R}^{n \times m}; \quad Y = \begin{bmatrix} x_1 \ldots x_m \end{bmatrix} \in \mathbb{R}^{n \times m}. \tag{3.87}$$

Notice that the matrix $Y$ is shifted forward compared to the matrix $X$ by $\Delta t$. Next, we solve for a matrix $B$ that minimizes the distance between $Y$ and $BX$. That is, we find

$$B_* = \underset{B \in \mathbb{R}^{n \times n}}{\operatorname{argmin}}\, g(B); \quad g(B) = \|Y - BX\|^2, \tag{3.88}$$

where the matrix norm is the Frobenius norm. We denote the components of $B$ by $b_{ij}$, the

components of $Y$ by $y_{ij}$, and the components of $X$ by $x_{ij}$. Note that

$$\frac{\partial g}{\partial b_{ij}} = \frac{\partial}{\partial b_{ij}} \sum_{l,m,p} (y_{lm} - b_{lp}x_{pm})^2 = -2 \sum_{l,m,p} (y_{lm} - b_{lp}x_{pm})\, \delta_{il}\delta_{jp}x_{pm} \tag{3.89a}$$

$$= -2 \sum_{m,p} (y_{im} - b_{ip}x_{pm})\, x_{jm} = 2\left(YX^T - BXX^T\right). \tag{3.89b}$$

In the third equality we have used the properties of the Kronecker-delta $\delta_{ij}$. Therefore we must have that $D_B g = -2\left(YX^T - BXX^T\right) = 0$, so we must solve the linear system of equations

$$BXX^T = YX^T \in \mathbb{R}^{n\times n}. \tag{3.90}$$

In the case $n > m$, the matrix $X^T$ has more columns than rows, thus $X^T$ is singular. Hence $XX^T$ has a nonempty kernel and is therefore not invertible. For $n < m$ this is not necessarily the case. For a solution, we can use the pseudo inverse $(XX^T)^\dagger$, the superscript here is called *dagger*.

Recall that if $A \in \mathbb{R}^{n\times n}$ is invertible then $A^{-1} = A^\dagger$. If $A$ is not invertible, then calculate the singular value decomposition of $A = U\Sigma V^T$ for $U, V \in \mathrm{SO}(n)$ and $\Sigma$ a diagonal matrix. The entries of $\Sigma$ are the strictly positive singular values. The columns of $U$ and $V$ are the left and right singular vectors. We then define the pseudo inverse as

$$A^\dagger = V\Sigma^\dagger U^T; \quad \Sigma^\dagger_{ij} = \begin{cases} \frac{1}{\Sigma_{ij}} & \Sigma_{ij} \neq 0 \\ 0 & \Sigma_{ij} = 0. \end{cases} \tag{3.91}$$

This means that for $Ax = b$, the best solution in a least squares sense is given by $x = A^\dagger b$. In our case, the solution to 3.88 is

$$B = \left(YX^T\right)\left(XX^T\right)^\dagger. \tag{3.92}$$

This procedure is called *dynamic mode decomposition* (DMD) [Schmid, 2010, Kutz et al., 2015]. The eigenvectors of $B$ are called *DMD modes* and the normalized exponents of those corresponding eigenvalues are called *DMD eigenvalues*.

The procedure has been extended in *extended DMD*, where DMD was performed on functions of the observables (typically polynomials) [Williams et al., 2015]. This effectively tries to construct the linearization from data. Note here that none of these procedures can describe intrinsically nonlinear phenomena (i.e. those with coexisting isolated compact invariant sets).

*Remark* 3.19. The numerical cost of (pseudo)inverting the $n \times n$ matrix, $XX^T$, can be reduced when $n \gg m$.

(i) We can calculate the singular value decomposition of the data matrix $X$ instead.

(ii) By keeping only the leading $r$ singular values in this decomposition, we obtain a reduced (rank-$r$) approximation of the matrix $B \in \mathbb{R}^{n\times n}$ as follows. Denote the singular value decomposition by $X = U\Sigma V^T$. This also means that $(XX^T) = U\Sigma^2 U^T$, i. e. we get the same $U$ as before, but from an SVD of an $n \times m$ matrix, which is much less demanding numerically.

With the leading $r$ singular vectors and singular values $\sigma$, $X$ can be written as

$$X \approx U_r \Sigma_r V_r^T = \sum_{i=1}^{r} \sigma_i u_i v_i^T, \qquad U = [u_1, ..., u_n] \quad V = [v_1, ..., v_m]. \tag{3.93}$$

(iii) The matrix $B$ can be approximately written as

$$B = YX^T(XX^T)^\dagger \tag{3.94a}$$
$$\approx YV_r\Sigma_r U_r^T U_r(\Sigma_r^2)^\dagger U_r^T \tag{3.94b}$$
$$= YV_r\Sigma_r^\dagger U_r^T = Y(X_r)^\dagger. \tag{3.94c}$$

We denote the truncated approximation of $B$ as $B_r = Y(X_r)^\dagger$.

**Exercise 3.4.** *Consider the damped pendulum equation*

$$\ddot{x} + c\dot{x} + \sin(x) = 0. \tag{3.95}$$

*Using the energy of the pendulum as a Lyapunov function, can we conclude the nonlinear asymptotic stability of the $x = 0$ equilibrium? Give detailed reasoning why.*

## 3.7 Lyapunov's direct (second) method for stability

Stability analysis via linearization is not perfect. For nonhyperbolic fixed points, the results obtained are inconclusive, and this case turns out to be somewhat common for conservative systems. Furthermore, linearization does not give any insight into the size of the domain of stability. Hence, a method for determining stability type without relying on linearization would be desirable. This is given to us by Lyapunov's direct method.

**Theorem 3.20.** *Consider*

$$\dot{x} = f(x), \quad f \in \mathcal{C}^1, \quad x \in \mathbb{R}^n; \quad f(x_0) = 0. \tag{3.96}$$

*Assume that there exists a function $V : U \to \mathbb{R}$ with $V \in \mathcal{C}^1(U)$ for an open set $U \subset \mathbb{R}^n$ and $x_0 \in U$ which fulfills the following*

*(i) $V$ is positive definite in a neighborhood of $x_0$, i.e.*

$$V(x_0) = 0; \quad V(x) > 0, \ x \in U - \{x_0\}. \tag{3.97}$$

*(ii) $\dot{V}$ is negative semidefinite in the same neighborhood, i.e.*

$$\dot{V} = \frac{d}{dt}V(x(t)) = \langle DV(x(t)), \dot{x}(t)\rangle = \langle DV(x(t)), f(x(t))\rangle \leq 0, \ x \in U. \tag{3.98}$$

*Then $x = x_0$ is Lyapunov stable (cf. Def. 3.1). $V$ is called a Lyapunov function. The hypotheses are illustrated in Fig. 3.21.*

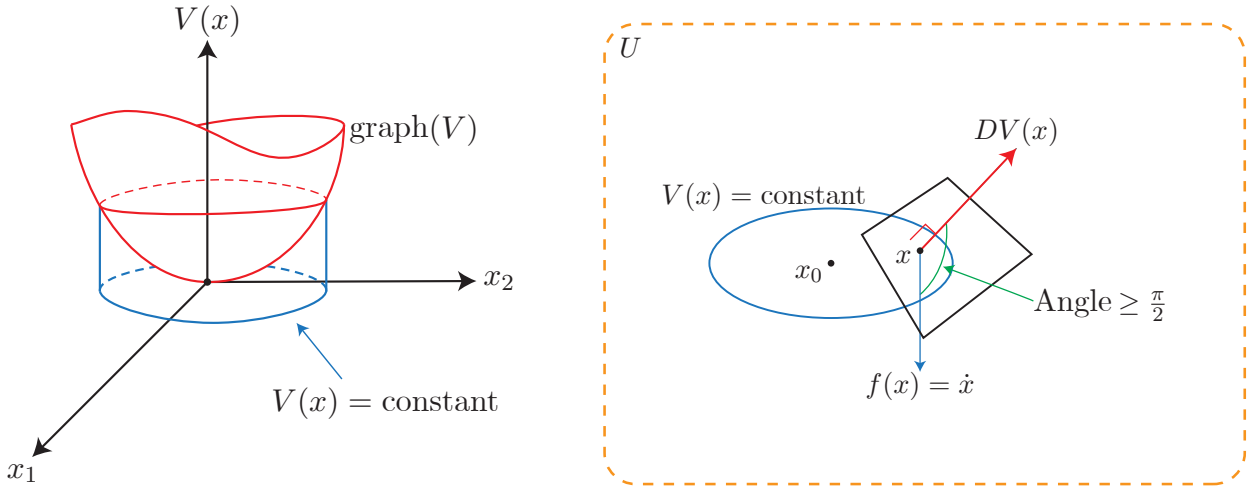


Figure 3.21: Geometric interpretation of hypotheses of Lyapunov's direct method. Left depicts hypothesis (i) and right hypothesis (ii). The angle between $DV(x)$ and $f(x)$ is at least $\frac{\pi}{2}$. In each image the blue $V(x)$ =const. is a level surface diffeomorphic to $S^{n-1}$.

*Remark* 3.21. To denote the boundary of a set $A$ we write $\partial A$.

*Proof.* First choose $\varepsilon > 0$ and define $\alpha(\varepsilon) = \min_{x \in \partial B_\varepsilon(x_0)} V(x) > 0$. Note that $\alpha(\varepsilon)$ is well defined as $V \in \mathcal{C}^0(U)$ and $\partial B_\varepsilon(x)$ is compact and spherical for small enough $\varepsilon$. There exists an $x^* \in \partial B_\varepsilon(x)$ with $V(x^*) = \alpha(\varepsilon) \leq V(x)$ for all $x \in \partial B_\varepsilon(x_0)$. Next, define $U_\varepsilon = \{x \in B_\varepsilon(x_0) : V(x) < \alpha(\varepsilon)$ . Notice that $x_0 \in U_\varepsilon$ because $V(x_0) = 0$ and $V(x) \geq 0$ on $U$. Further $U_\varepsilon$ is open due to the continuity of $V$. We have that $U_\varepsilon \cap \partial B_\varepsilon(x_0)$ is empty by definintion, noting that for all $x \in \partial B_\varepsilon(x_0)$ we have $V(x) \geq \alpha(\varepsilon)$. Therefore there exists a ball $B_{\delta(\varepsilon)} \subset U_\varepsilon$ which contains $x_0$. This can be seen in Fig. 3.22.

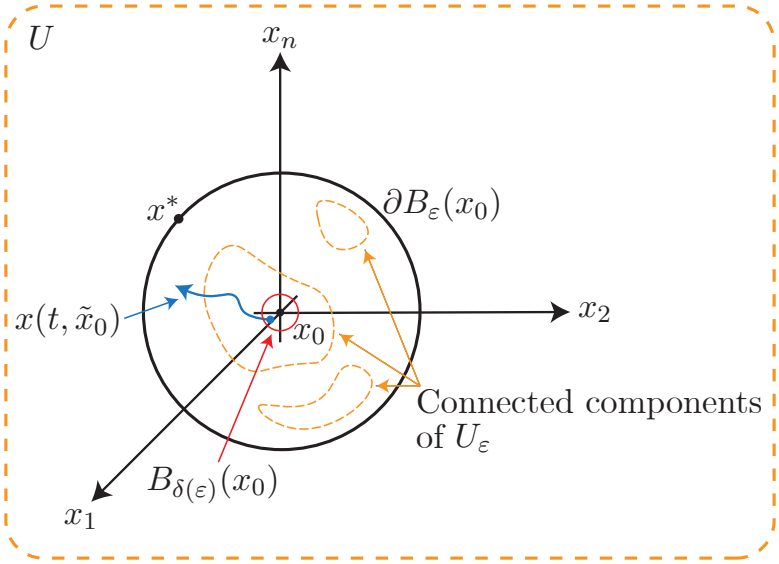


Figure 3.22: The constellation of $U$, $U_\varepsilon$, $\partial B_\varepsilon(x_0)$, and $B_{\delta(\varepsilon)}(x_0)$ from the proof of Lyapunov's direct method. The red arrow points at $B_{\delta(\varepsilon)}(x_0)$ and the yellow arrow at a connected component of $U_\varepsilon$.

Now observe that for every $\tilde{x}_0 \in B_{\delta(\varepsilon)}(x_0)$ we have that along trajectories $V(x(t;\tilde{x}_0)) \leq V(\tilde{x}_0) < \alpha(\varepsilon)$. The first inequality comes from hypothesis (ii) and the second inequality from the definition of $U_\varepsilon$. This implies that for $x(t;\tilde{x}_0) \in U_\varepsilon$ we have that $x(t;\tilde{x}_0)$ is not in $\partial B_\varepsilon(x_0)$ for any $t > 0$. The trajectory $x(t;\tilde{x}_0)$ is continuous, in order for it to leave the ball $B_\varepsilon(x_0)$ it must intersect the boundary $\partial B_\varepsilon(x_0)$. At this point $V(x)$ will atain a value of at least $\alpha(\varepsilon)$ by definition, however this is in contradiction to the fact that $V(x(t;\tilde{x}_0))$ is strictly smaller that $\alpha(\varepsilon)$. Therefore must stay in the ball $B_\varepsilon(x_0)$ for all times. □

Now we present some extensions to this theorem for various different stability types. These are necessary, as the previous theorem only provides a sufficient condition for stability (it is not "if and only if").

**Theorem 3.22** (Theorem 2)**.** *Consider the same dynamical system. Assume*

*(i)* $V(x)$ *is positive definite,*

*(ii)* $\dot{V}(x)$ *is negative definite, i.e.*

$$\dot{V}(x) < 0, \ x \in U - \{x_0\}. \tag{3.99}$$

*Then* $x = x_0$ *is asymptotically stable. These hypotheses are illustrated in Fig. 3.23.*

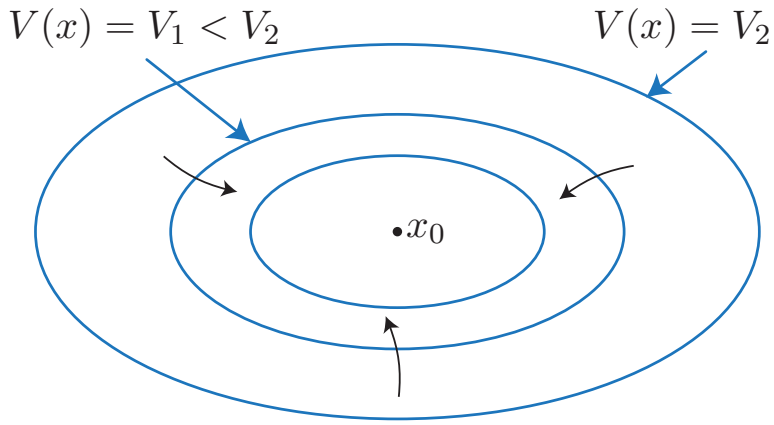


Figure 3.23: The hypotheses of Theorem 2 illustrated, the key difference being that the arrows (denoting the flow of the dynamical system) cross the level surfaces of $V$ (blue rings) toward $x_0$.

**Theorem 3.23** (Theorem 3)**.** *Consider the same dynamical system. Assume*

*(i)* $V(x)$ *is positive definite,*

*(ii)* $\dot{V}(x)$ *is positive definite, i.e.*

$$\dot{V}(x) > 0, \ x \in U - \{x_0\}. \tag{3.100}$$

*Then* $x = x_0$ *is unstable. The hypotheses are illustrated in Fig. 3.24.*

**Theorem 3.24** (Theorem 4)**.** *Consider the same dynamical system. Assume*

*(i)* $V(x)$ *is indefinite, i.e. arbitrarily close to* $x_0$ *there exists* $a, b \in U$ *such that* $V(x_1) \cdot V(x_2) < 0$ *(they have opposite signs and are not equal to 0) and* $V(x_0) = 0$*. Furthermore, this is a critical point, i.e.* $DV(x_0) = 0$*.*

*(ii)* $\dot{V}(x)$ *is definite near* $x_0$ *(either positive or negative).*

*Then* $x = x_0$ *is unstable. The geometry of the hypotheses are illustrated in Fig. 3.25.*

*Remark* 3.25. In each of these theorems, the definiteness of $\dot{V}$ can be replaced by semidefiniteness, if we add that the set $\{x \in U : \dot{V}(x) = 0\}$ does not contain full trajectories of the system. This is called Krasovsky's condition.

Now we would like to put these theorems into practice with a few examples.

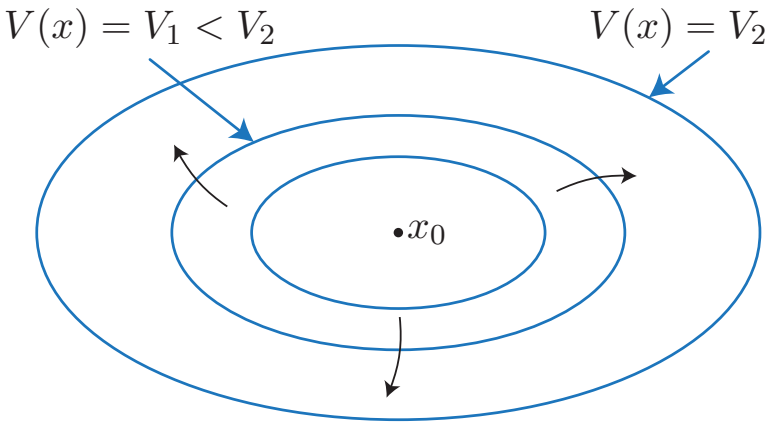


Figure 3.24: The hypotheses of Theorem 3 illustrated, the key difference being that the arrows (denoting the flow of the dynamical system) cross the level surfaces of $V$ (blue rings) away from $x_0$.

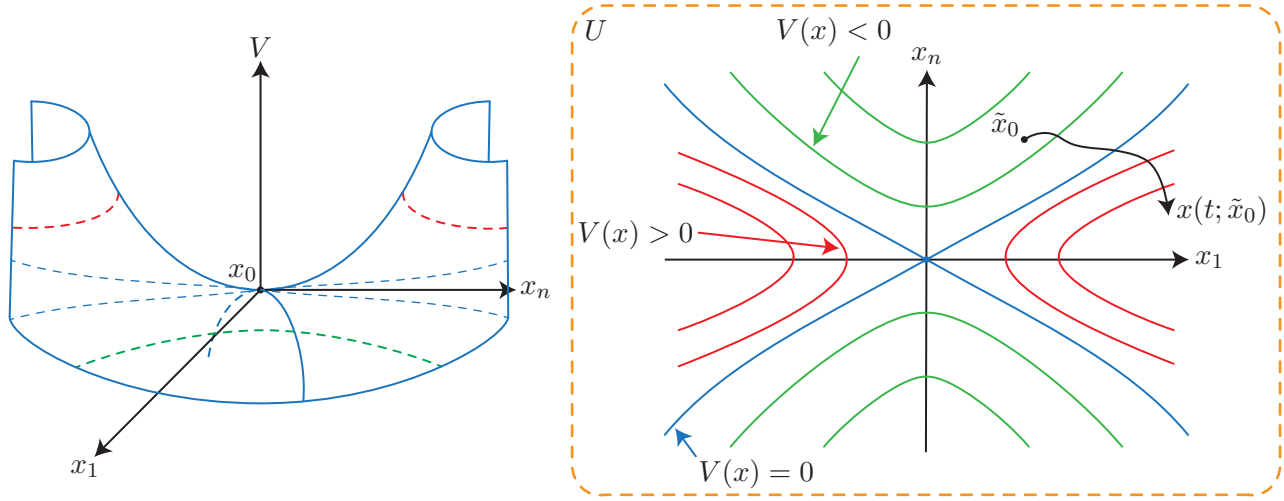


Figure 3.25: The geometry of hypotheses (i) (left) and (ii) (right) for $\dot{V}$ positive definite of Theorem 4 illustrated. On the right level surfaces are designated by lines, blue corresponds to $V = 0$, red $V > 0$, and green $V < 0$, which can be seen as the dotted lines of the same colors on the left.

*Example* 3.12 (Stability analysis of the pendulum with Lyapunov's direct method). Recall that using linearization, we were only able to conclude the stability type of one of the fixed points for the dynamical system of the pendulum

$$ml^2\ddot{\varphi} + mgl\sin(\varphi) = 0. \tag{3.101}$$

Now we will use the energy as a Lyapunov function, which is often very useful. The energy is given by

$$E(x) = E(\varphi, \dot{\varphi}) = \frac{1}{2} ml^2 \dot{\varphi}^2 + mgl(1 - \cos(\varphi)) = \frac{1}{2} ml^2 \dot{x_2}^2 + mgl(1 - \cos(x_1)). \tag{3.102}$$

Transforming the dynamical system to be an system of first order ODEs we obtain

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \varphi \\ \dot{\varphi} \end{pmatrix}; \quad \dot{x} = \begin{pmatrix} \dot{x_1} \\ \dot{x_2} \end{pmatrix} = \begin{pmatrix} x_2 \\ -\frac{g}{l} \sin(x_1) \end{pmatrix} = f(x). \tag{3.103}$$

At the fixed point $x = (0, 0)$ we have

$$E(0,0) = 0; \quad DE(0,0) = 0 \in \mathbb{R}^{2\times 2}; \quad D^2E(0,0) = \begin{pmatrix} mgl & 0 \\ 0 & ml^2 \end{pmatrix}. \tag{3.104}$$

We have that the Hessian of $E$ is positive definite. Therefore $E$ is positive definite at $(0, 0)$. Further we have

$$\dot{E}(x) = \langle DE(x), f(x) \rangle = \begin{pmatrix} mgl \sin(x_1) & ml^2 x_2 \end{pmatrix} \begin{pmatrix} x_2 \\ -\frac{g}{l} \sin(x_1) \end{pmatrix} = 0. \tag{3.105}$$

Thus $\dot{E}$ is negative semidefinite. Now the hypotheses of Theorem 3.20 are fulfilled, hence $x = (0, 0)$ is Lyapunov stable. Importantly, this is a nonlinear result and we did not need to refer to the linearization of the system!

At the fixed point $x = (\pi, 0)$. We check again using the energy. First we realize that $E(\pi, 0) = 2mgl$, so we subtract the constant and redefine to obtain $\tilde{E} = E - 2mgl$. Now $\tilde{E}(\pi, 0) = 0$, although this is not essential. Next we calculate

$$DE(\pi,0) = 0 \in \mathbb{R}^{n\times n}; \quad D^2E(\pi,0) = \begin{pmatrix} -mgl & 0 \\ 0 & ml^2 \end{pmatrix}. \tag{3.106}$$

Hence, $E$ is indefinite at $(\pi, 0)$. From the previous calcuation we alread know that $\dot{E}(\pi, 0) = 0$, i.e. $\dot{E}$ is semidefinite. We cannot apply Theorem 4, but linearization already concluded that $(\pi, 0)$ was unstable.

If we add damping the dynamics becomes

$$ml^2\ddot{\varphi} + c\dot{\varphi}mg\ell \sin(\varphi) = 0. \tag{3.107}$$

Again we transform the system by introducing the coordinates $x_1 = \varphi$ and $x_2 = \dot{\varphi}$. Thus we have a first order ODE

$$\begin{cases} \dot{x_1} = x_2 \\ \dot{x_2} = -\dfrac{c}{m\ell^2} x_2 - \dfrac{g}{\ell} \sin(x_1). \end{cases} \tag{3.108}$$

We still have the fixed points $(0,0)$ and $(0,\pi)$ and the energy $E = \frac{1}{2}m\ell^2\dot{\varphi}^2 + mg\ell(1-\cos(\varphi))$. Again we find

$$DE = \begin{pmatrix} mg\ell\sin(x_1) \\ m\ell^2 x_2 \end{pmatrix} \implies DE(0) = 0; \quad D^2E = \begin{pmatrix} mg\ell\cos(x_1) & 0 \\ 0 & m\ell^2 \end{pmatrix}. \tag{3.109}$$

Thus $E$ is positive definite at $(0,0)$ and indefinite at $(0,\pi)$. We have

$$\dot{E} = DE \cdot f = mgl\sin(x_1)x_2 + m\ell^2 x_2\left(-\frac{c}{m\ell^2}x^2 - \frac{g}{\ell}\sin(x_1)\right) = -cx_2^2. \tag{3.110}$$

Note here that $\dot{E} = -cx_2^2$ is only negative semi-definite. Therefore the asymptotic stability of $x = (0,0)$ and the instability of $x = (0,\pi)$ do not follow from theorems 1 or 4. However, the set $S = \left\{x : \ \dot{E}(x) = 0\right\} - \{(0,0),\ (0,\pi)\}$ does <u>not</u> contain full trajectories. Indeed

$$\left.\frac{d}{dt}x_2(t)\right|_{x=0} = -\frac{g}{\ell}\sin(x_1) \neq 0, \ x \in S. \tag{3.111}$$

Hence, the asymptotic stability of $(0,0)$ and instability of $(0,\pi)$ follows by Krasovsky's condition.

*Example* 3.13 (Stability analysis of the friction pendulum). We have a shaft which is constantly rotating with angular speed $\Omega$, around this shaft is a sleeve which rubs against the shaft creating friction. To this sleeve is a mass $m$ attached at distance $l$, the deflection of this mass from its standard position (directly below the shaft) is measured by $\varphi$. The gravity constant is given by $g$. This setup is depicted in Fig. 3.26. The torque driving the pendulum is given by

$$T(\dot{\varphi}) = T_0\text{sign}(\Omega - \dot{\varphi}). \tag{3.112}$$

From this we get the equation of motion

$$ml^2\ddot{\varphi} + mgl\sin(\varphi) = T(\dot{\varphi}) = T_0\text{sign}(\Omega - \dot{\varphi}). \tag{3.113}$$

By assuming $\Omega \gg 1$, i.e. very fast rotation of the shaft, we find that $T(\dot{\varphi}) = T_0$. Next we transform the coordinates in order to get an ODE

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \varphi \\ \dot{\varphi} \end{pmatrix}; \quad \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ -\frac{g}{l}\sin(x_1) + \frac{T_0}{ml^2} \end{pmatrix}. \tag{3.114}$$

Thus the fixed point $x_0$ is at $(\overline{x}_1, \overline{x}_2)$ given by

$$\sin(\overline{x}_1) = \frac{T_0}{mgl}; \quad \overline{x}_2 = 0. \tag{3.115}$$

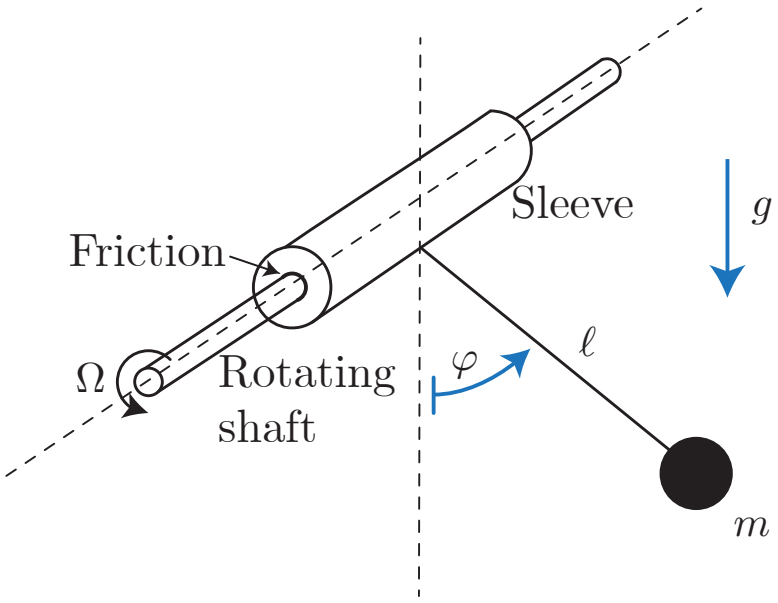


Figure 3.26: The setup of the friction pendulum.

Since the system is forced, energy is not conserved and we cannot use it as a Lyapunov function. However, there may still exist a conserved quantity. Consider using

$$V(x) = [\text{total energy at time } t] - [\text{work put in between time } t \text{ and } t_0] = [\text{initial energy}] = \text{const.} \tag{3.116}$$

This is a conserved quantity in all mechanics problems, but generally cannot be calculated without already knowing the trajectories. Here we can, as

$$V(x(t)) = \frac{1}{2}mlx_2^2 + mgl(1-\cos(x_1)) - T_0(\underbrace{x_1}_{\varphi(t)} - \underbrace{x_1(0)}_{\varphi(0)}). \tag{3.117}$$

We can drop the constant term $T_0 x_1(0)$. Now verify that $V$ is indeed constant

$$\dot{V}(x) = \frac{\partial V}{\partial x_1}\dot{x}_1 + \frac{\partial V}{\partial x_2}\dot{x}_2 = (mgl\sin(x_1) - T_0)\,x_2 + ml^2 x_2\left(-\frac{g}{l}\sin(x_1) + \frac{T_0}{ml^2}\right) = 0. \tag{3.118}$$

Hence $\dot{V}(x)$ is semidefinite, and may be used to conclude stability or instability. We must check the Lyapunov conditions for $V(x)$

$$DV(x_0) = \begin{pmatrix} mgl\sin(\overline{x}_1) - T_0 & ml^2\overline{x}_2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \end{pmatrix} \tag{3.119a}$$

$$D^2V(x_0) = \begin{pmatrix} mgl\cos(\overline{x}_1) & 0 \\ 0 & ml^2 \end{pmatrix}. \tag{3.119b}$$

The Hessian is positive definite as long as $\overline{x}_1 \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$, and all fixed points in this region are stable by Theorem 3.20. The Hessian is indefinite if $\overline{x}_1 \notin \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, but in this case Theorem 4 is not applicable as $\dot{V}$ is not definite. More detailed discussion and further results on Lyapunov's direct method can be found in [Rouche et al., 1977].

**Exercise 3.5.** *Recall from Exercise 3.2 that a ball of mass $m$ sliding on a hoop rotating with angular velocity $\Omega$ satisfies the differential equation*

$$mR^2\ddot{\alpha} + mR^2(g/R - \Omega^2\cos(\alpha))\sin(\alpha) = 0 \tag{3.120}$$

*if there is no friction between the hoop and the mass, i.e. we assume $b = 0$. Assume that the parameter values are such that the lower equilibrium position of the ball is stable in linear approximation.*

*(a) Show that in this case, the equilibrium is also nonlinearly stable.*

*Hint: Note that system (3.120) is not conservative: an external torque is required to maintain the constant rotation of the hoop. As a result, the total mechanical energy of the system is not expected to work as a Lyapunov function. To find another candidate for a Lyapunov function, find a quantity that is conserved along trajectories. Such a quantity can be found by multiplying equation (3.120) by $\dot{\alpha}$ and integrating in time.*

*(b) Prove that the equilibrium cannot be asymptotically stable for the nonlinear system.*

*Hint: use the Lyapunov function you have found in (a).*

**Exercise 3.6.** *Consider the damped pendulum equation*

$$\ddot{x} + c\dot{x} + \sin(x) = 0. \tag{3.121}$$

*(a) A theorem due to Krasovski states the following: Assume that $x = 0$ is a fixed point for the $n$-dimensional dynamical system $\dot{x} = f(x)$. Assume that there exists a smooth scalar function $V(x)$ such that*

*(i) $V(x)$ is positive definite on an open neighborhood $U$ of $x = 0$;*

*(ii) $\dot{V}$ is negative semi-definite on the same neighborhood;*

*(iii) the only trajectory lying completely in the set $S = \{x \in U : \dot{V} = 0\}$ is the fixed point $x = 0$. Then $x = 0$ is asymptotically stable.*

*Use Krasovski's theorem to conclude the asymptotic stability of the origin for (3.121).*

**Exercise 3.7.** *Consider an n-degree-of-freedom holonomic mechanical system (i.e. one that has only position-dependent constraints) with generalized coordinates $q \in \mathbb{R}^n$ and generalized velocities $\dot{q} \in \mathbb{R}^n$. The total energy of the system is of the form*

$$E(q, \dot{q}) = \frac{1}{2}\dot{q}^T M(q)\dot{q} + V(q), \tag{3.122}$$

*where $M \in \mathbb{R}^{n\times n}$ is the mass matrix (symmetric ans positive definite), and $V(q)$ is the potential energy. In the absence of external forces, the associated Lagrangian equations of motion are*

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0, \tag{3.123}$$

*where $L(q, \dot{q}) = \frac{1}{2}\dot{q}^T M(q)\dot{q} - V(q)$ is the Lagrangian of the mechanical system.*

*Show that if $V(q)$ admits a strict local minimum at a point $q_0$, then $q_0$ is a (nonlinearly) stable equilibrium for the mechanical system. (This result is also known as Dirichlet's Theorem in classical mechanics).*

# Chapter 4

# Bifurcations of fixed points

## 4.1 Local nonlinear dynamics near fixed points

We are interested in the local nonlinear dynamics around fixed points. Consider

$$\dot{x} = f(x); \quad f \in \mathcal{C}^r, r \geq 1; \quad f(p) = 0, \tag{4.1}$$

i.e. $p$ is a fixed point of the dynamical system. The linearized system at $p$ is

$$\dot{y} = Df(p)y, \ y \in \mathbb{R}^n, \ Df(p) \in \mathbb{R}^{n \times n}. \tag{4.2}$$

The linearization has the eigenvalues $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ with multiplicities counted. Corresponding to these eigenvalues are the eigenvectors $e_1, \ldots, e_n \in \mathbb{C}^n$, including generalized eigenvectors for when the algebraic multiplicity is greater than the geometric multiplicity. The eigenvector $e_j$ is real when $\lambda_j \in \mathbb{R}$.

**Definition 4.1.** The following subspaces are invariant for the linearized dynamical system:

(i) The *stable subspace*

$$\boxed{E^S = \text{span}_j \left\{ \text{Re}(e_j), \text{Im}(e_j) : \ \text{Re}(\lambda_j) < 0 \right\},} \tag{4.3}$$

(ii) The *unstable subspace*

$$\boxed{E^U = \text{span}_j \left\{ \text{Re}(e_j), \text{Im}(e_j) : \ \text{Re}(\lambda_j) > 0 \right\},} \tag{4.4}$$

(iii) The *center subspace*

$$\boxed{E^C = \text{span}_j \left\{ \text{Re}(e_j), \text{Im}(e_j) : \ \text{Re}(\lambda_j) = 0 \right\}.} \tag{4.5}$$

*Remark* 4.1. Note here that the following facts hold

(i) $E^C = \emptyset$ if and only if $p$ is hyperbolic,

(ii) $E^{U,S}$ and $E^C$ are invariant subspaces of (4.2) by construction,

(iii) Solutions of (4.2) in $E^S$ (resp. $E^U$) decay to $y = 0$ as $t \to \infty$ (resp. $t \to -\infty$).

We now discuss what happens to these subspaces in the nonlinear system (4.1).

**Theorem 4.2** (Center Manifold Theorem)**.** *The following hold:*

*(i) There exists a unique* stable manifold $W^S(p)$ *for* (4.1)*, such that*

- $W^S(p)$ *is a* $\mathcal{C}^r$ *manifold (surface), tangent to* $E^S$ *at* $p$ *with dim* $W^S(p) =$ *dim* $E^S$*,*
- $W^S(p)$ *is invariant for* (4.1) *and for* $x \in W^S(p)$ *we have*

$$\|F^t(x) - p\| \leq K_S \exp\left[t\left(\max_{Re(\lambda_j)<0}(Re(\lambda_j)) + \varepsilon_S\right)\right] \tag{4.6}$$

*for* $t \geq 0$*,* $0 < \varepsilon_S \ll 1$*, and* $\|x - p\|$ *small enough.*

*(ii) There exists a unique* unstable manifold $W^U(p)$ *for* (4.1)*, such that*

- $W^U(p)$ *is a* $\mathcal{C}^r$ *manifold (surface), tangent to* $E^U$ *at* $p$ *with dim* $W^U(p) =$ *dim* $E^U$*,*
- $W^U(p)$ *is invariant for* (4.1) *and for* $x \in W^U(p)$ *we have*

$$\|F^t(x) - p\| \leq K_U \exp\left[t\left(\min_{Re(\lambda_j)>0}(Re(\lambda_j)) - \varepsilon_U\right)\right] \tag{4.7}$$

*for* $t \leq 0$*,* $0 < \varepsilon_U \ll 1$*, and* $\|x - p\|$ *small enough.*

*(iii) There exists a (not necessarily unique)* center manifold $W^C(p)$ *for* (4.1)*, such that*

- $W^C(p)$ *is a* $\mathcal{C}^{r-1}$ *manifold (surface), tangent to* $E^C$ *at* $p$ *with dim* $W^C(p) =$ *dim* $E^C$*,*

The geometry of these manifolds and their corresponding subspaces are sketched in Fig. 4.1.

Notice that Theorem 4.2 does not state anything about the asymptotic dynamics on the center manifold, in contrast to that of the unstable and stable manifolds. This means that the overall dynamics depends crucially on the center manifold. In the case when $W^u = \emptyset$, even the stability type of the fixed point is determined by $W^C(p)$. This is why it will be the subject of further investigations.

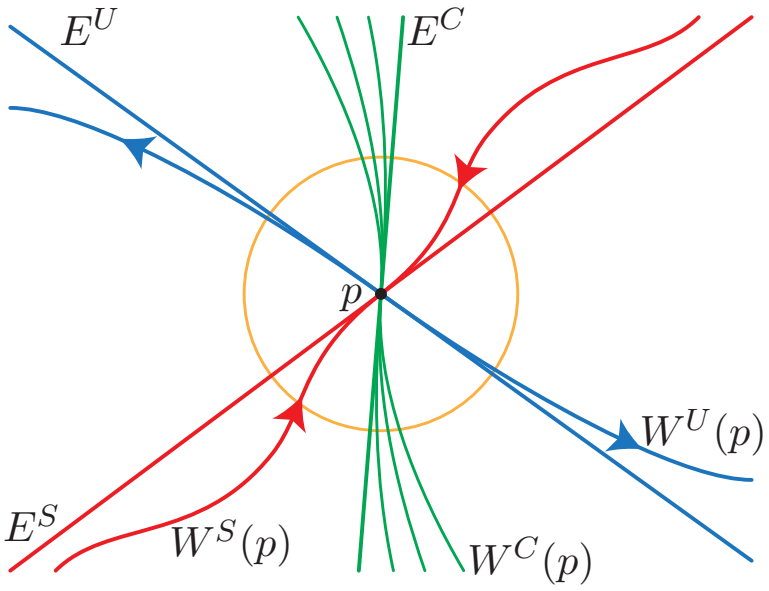


Figure 4.1: A sketch of the stable (red), unstable (blue), and center manifolds (green), along with their respective linear subspaces, that are invariant under the linearized dynamics. Note the existence of multiple center manifolds and the unique unstable/stable manifolds.

## 4.2 The center manifold

*Example* 4.1 (Uniqueness of the center manifold). Since Theorem 4.2 gave no guarantees on the uniqueness of the center manifold, now we explore this non-uniqueness in an example.

$$\begin{cases} \dot{x} = x^2 \\ \dot{y} = -y. \end{cases} \tag{4.8}$$

First, linearize at the origin to find the linearized dynamics

$$A = Df(0) = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}. \tag{4.9}$$

The linearized dynamics is illustrated in Fig. 4.2. We find the invariant subspaces

$$E^C = \text{span}\left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right\}; \quad E^S = \text{span}\left\{\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right\}; \quad E^U = \emptyset. \tag{4.10}$$

The nonlinear manifolds are illustrated with the invariant subspaces from the linearization in Fig. 4.2. Observe that there exist infinitely many center manifolds which are all invariant and all tangent to $E^C$ at the origin. We also see that although the fixed point at the origin is stable in the linearized system, it is unstable in the nonlinear system.

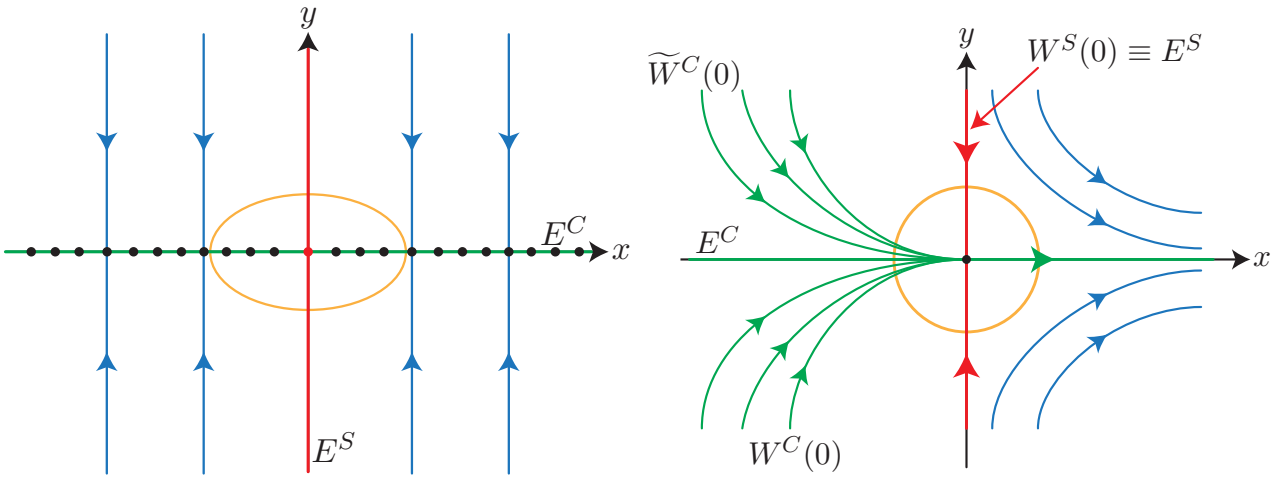


Figure 4.2: Left: The linearized dynamics around the origin. Right: The nonlinear phase portrait.

*Remark* 4.3. All center manifolds $W^C(0)$ are $\mathcal{C}^\infty$, but not analytic at $x = 0$, for instance

$$y(x) = y_0 e^{-\frac{1}{x_0}} e^{\frac{1}{x}} \implies \forall n \in \mathbb{N} : \frac{d^n}{dx^n} y(0) = 0. \tag{4.11}$$

We now present a method to calculate $W^C(p)$ in general cases.

(i) Consider

$$\dot{z} = F(z); \quad F(0) = 0; \quad z \in \mathbb{R}^{c+d}; \quad F \in \mathcal{C}^r, \tag{4.12}$$

where $c$ represents the number of center directions at the origin (dim $E^C$) and $d$ denotes the remaining directions (dim $E^U$ + dim $E^S$).

(ii) Now block-diagonalize the linearization. This consists of four steps

(a) First linearize the dynamics to find $\dot{z} = Mz + \mathcal{O}(\|z\|^2)$ with $M = DF(0) \in \mathbb{R}^{(c+d)\times(c+d)}$.

(b) Define the transformation matrix

$$T = \begin{bmatrix} a_1 & \dots & a_c & b_1 & \dots & b_d \end{bmatrix} = \begin{bmatrix} \text{basis in } E^C & \text{basis in } E^U \oplus E^S \end{bmatrix}. \tag{4.13}$$

(c) Pass to the basis from the transformation matrix by introducing a new set of coordinates as $z = T\xi$

$$\dot{\xi} = T^{-1}\dot{z} = T^{-1}MT\xi + T^{-1}\mathcal{O}(\|T\xi\|^2) = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \xi + \mathcal{O}(\|\xi\|^2). \tag{4.14}$$

The matrices $A$ and $B$ are elements of $\mathbb{R}^{c\times c}$ and $\mathbb{R}^{d\times d}$ respectively.

(d) Let $\xi = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^c \times \mathbb{R}^d$, the $x$-coordinates are aligned with $E^C$ and the $y$-coordinates are perpendicular to them. We then find

$$\dot{x} = Ax + f(x, y); \quad \dot{y} = By + g(x, y), \tag{4.15}$$

for $f, g \in \mathcal{C}^r$ and $f, g = \mathcal{O}(\|x\|^2, \|y\|^2, \|x\|\|y\|)$. The geometry in these coordinates is depicted in Fig. 4.3. The center manifold is given by

$$W^C(0) = \{(x, y) \in U : \ y = h(x)\} \tag{4.16}$$

for $h : \mathbb{R}^c \to \mathbb{R}^d$ and $h \in \mathcal{C}^{r-1}$ as in the theorem.

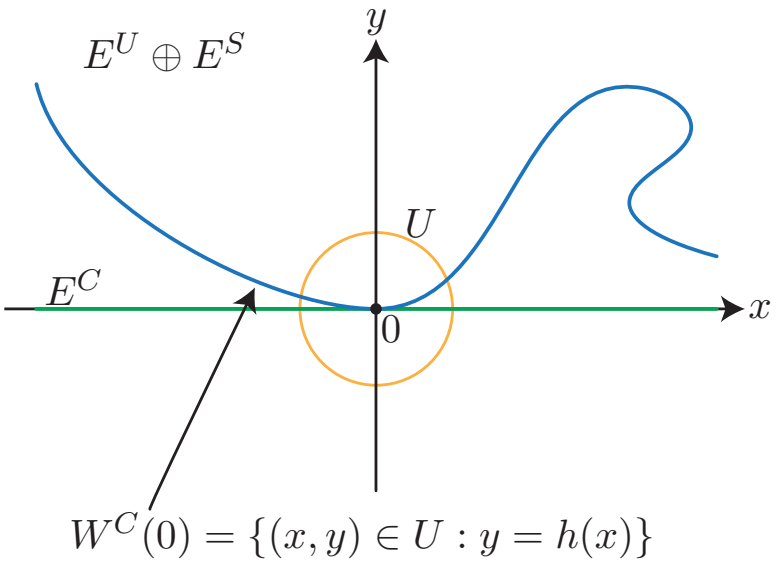


Figure 4.3: The geometry of the nonlinear system in the transformed coordinates aligned with the invariant subspaces of the linearization.

(iii) Now we use the invariance of $W^C(0)$, i.e. for all $t$, it holds that $y(t) = h(x(t))$. With this we find

$$\dot{y} = Dh(x(t))\dot{x}(t). \tag{4.17}$$

We can now substitute $\dot{x}$ and $\dot{y}$ from (4.15) to get the following nonlinear partial differential equation (PDE) for $h(x)$

$$\boxed{Bh(x) + g(x, h(x)) = Dh(x)\left[Ax + f(x, h(x))\right].} \tag{4.18}$$

In addition to (4.18) being nonlinear, the boundary conditions are also unknown. Therefore there is little hope for solving it analytically.

(iv) Instead take the Taylor expansion of (4.18) to approximate the solution

$$h(x) = \underbrace{h(0)}_{=0} + \underbrace{Dh(0)}_{=0} x + \frac{1}{2} \underbrace{D^2h(0)}_{3-\text{tensor}} \otimes x \otimes x + \mathcal{O}(\|x\|^3), \tag{4.19}$$

where the first two terms are 0 due to the tangency to $E^C$ at 0. We therefore have that $h = O(x^2)$ and we are justified in seeking $W^C(0)$ in this form. We can then restrict the system to the center manifold by substituting $y = h(x)$ to get the reduced dynamics as

$$\boxed{\dot{x} = Ax + f(x, h(x)).} \tag{4.20}$$

**Exercise 4.1.** *Consider the discrete dynamical system*

$$x_{n+1} = Ax_n + f(x_n, y_n), \tag{4.21}$$
$$y_{n+1} = By_n + g(x_n, y_n), \tag{4.22}$$

*where $x_n \in \mathbb{R}^c$, $y_n \in \mathbb{R}^d$, $A \in \mathbb{R}^{c\times c}$, $B \in \mathbb{R}^{d\times d}$; $f$ and $g$ are $\mathcal{C}^r$ functions with no linear terms. Assume that all eigenvalues of $A$ have modulus one, and none of the eigenvalues of $B$ have modulus one. Then the linearized system at the origin admits a center subspace $E^c$ aligned with the $x$-coordinate plane.*

*(a) Derive a general algebraic equation for the center manifold $W^C$, which is known to exist by a theorem analogous to the center manifold theorem for continuous dynamical systems.*

*(b) Find a cubic order approximation for the center manifold of the discrete system*

$$x_{n+1} = x_n + x_n y_n, \tag{4.23}$$
$$y_{n+1} = \lambda y_n - x_n^2, \tag{4.24}$$

*where $\lambda \in (0, 1)$.*

*(c) Reduce the dynamics to the center manifold and determine the stability of the origin. Verify your results by a numerical simulation of a few initial conditions near the origin.*

*Example* 4.2 (Finding the center manifold). Consider the dynamical system

$$\begin{cases} \dot{x} = xy \\ \dot{y} = -y + \alpha x^2. \end{cases} \tag{4.25}$$

First we linearize at $(0,0)$ to get

$$M = \begin{pmatrix} [0] & [0] \\ [0] & [-1] \end{pmatrix} \tag{4.26}$$

which is already in block-diagonal form. The dimensions of the stable, unstable, and center subspaces of the linearization are 1, 0, and 1 respectively. Hence the stability type depends on the dynamics on the center manifold $W^C(0)$. We now look for a parameterization of $W^C(0)$ in the form

$$h(x) = ax^2 + bx^3 + cx^4 + \mathcal{O}(x^5). \tag{4.27}$$

Since the right-hand side of the dynamical system was $\mathcal{C}^\infty$, this expansion could formally be continued to any order. However, the expansion will not converge as that would imply the uniqueness of the center manifold. In this example, we work with a finite, order 4 truncation. Now use the invariance (the PDE (4.18) we derived above) to find

$$\dot{y} = h'(x)\dot{x} = \left[2ax + 3bx^2 + 4cx^3 + \mathcal{O}(x^4)\right] x \left[ax^2 + bx^3 + cx^4 + \mathcal{O}(x^5)\right]. \tag{4.28}$$

On the other hand, from the dynamical system we know

$$\dot{y} = -h(x) + \alpha x^2 = (\alpha - a)x^2 - bx^3 - cx^4 + \mathcal{O}(x^5). \tag{4.29}$$

Equations (4.28) and (4.29) must hold for all $x$ small enough, therefore the coefficients of equal powers of $x$ in these two equations must match.

$$\mathcal{O}(x^2): \ \alpha = a \tag{4.30a}$$
$$\mathcal{O}(x^3): \ b = 0 \tag{4.30b}$$
$$\mathcal{O}(x^4): \ 2a^2 = -c. \tag{4.30c}$$

Therefore we find

$$\boxed{h(x) = \alpha x^2 - 2\alpha^2 x^4 + \mathcal{O}(x^5).} \tag{4.31}$$

Then the dynamics on $W^C(0)$ becomes

$$\boxed{\dot{x} = xh(x) = \alpha x^3(1 - 2\alpha x^2) + \mathcal{O}(x^6).} \tag{4.32}$$

The dynamics is depicted in Fig. 4.4. For $\alpha > 0$ the origin is unstable, meanwhile for $\alpha < 0$ the origin is asymptotically stable.

The full local stable manifold for $\alpha < 0$ is $\overline{W}^S(0) = U$ and it is of dimension 2. The difference between $\overline{W}^S(0)$ and $W^S(0)$ is that in general the decay rate in $\overline{W}^S(0) - W^S(0)$ is generally weaker than the rate guaranteed in the Center Manifold Theorem.

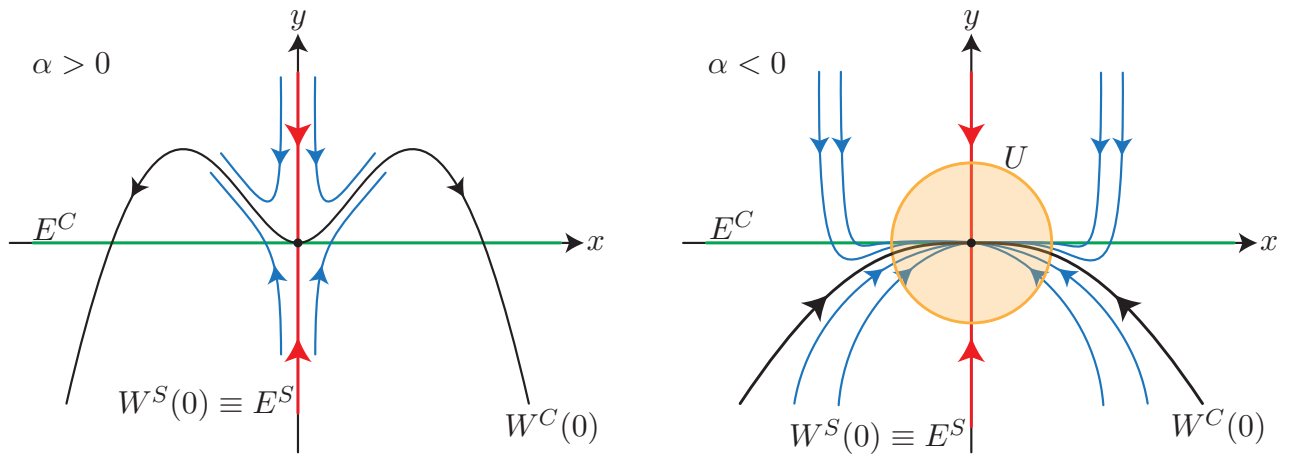


Figure 4.4: Left: The nonlinear dynamics on the center manifold for $\alpha > 0$. Right: The nonlinear dynamics on the center manifold for $\alpha < 0$.

*Remark* 4.4. The $\mathcal{O}(x^5)$ truncation has two hyperbolic fixed points at $x = \pm\frac{1}{\sqrt{2\alpha}}$, however the full system has no such fixed points. The reason for this is that away from the origin, the $\mathcal{O}(x^6)$ terms are no longer guaranteed to be small relative to the $\mathcal{O}(x^5)$ terms, and the truncation this far away from 0 is not justified.

After this example we would like to explore if the concept of the center manifold is robust, as the existence of eigenvalues with $\mathrm{Re}(\lambda_i) = 0$ is not. We will explore this in an example.

*Example* 4.3 (Perturbing the previous example). Consider the following perturbed dynamical system

$$\begin{cases} \dot{x} = xy + \varepsilon x \\ \dot{y} = -y + \alpha x^2 \end{cases}; \quad |\varepsilon| \ll 1. \tag{4.33}$$

The linearization of this system yields

$$\begin{cases} \dot{x} = \varepsilon x \\ \dot{y} = -y. \end{cases} \tag{4.34}$$

Now the center manifold disappears as the center subspace $E^C$ disappears for $\varepsilon \neq 0$!

## 4.3 Center manifolds depending on parameters

We begin with the setup

$$\begin{cases} \dot{x} = Ax + f(x, y, \varepsilon) \\ \dot{y} = By + g(x, y, \varepsilon) \end{cases}; \quad x \in \mathbb{R}^c,\ y \in \mathbb{R}^d,\ 0 \leq \|\varepsilon\| \ll 1;\ \varepsilon \in \mathbb{R}^p; \tag{4.35a}$$

$$f, g \in \mathcal{C}^r,\ f, g = \mathcal{O}(\|x\|^2, \|y\|^2, \|x\|\|y\|, \|\varepsilon\|\|x\|, \|\varepsilon\|\|y\|). \tag{4.35b}$$

The order $\|\varepsilon\|\|x\|$ and $\|\varepsilon\|\|y\|$ terms are due to the perturbation of the linear part. Now assume $\mathrm{Re}(\lambda_j(A)) = 0$ for $j = 1, \ldots, c$ (the center directions) and $\mathrm{Re}(\lambda_j(B)) \neq 0$ for $j = 1, \ldots, d$ (the hyperbolic directions). Next, rewrite $\tilde{x} = \begin{pmatrix} x \\ \varepsilon \end{pmatrix}$ and $\tilde{y} = y$ to obtain the system

$$\begin{cases} \dot{\tilde{x}} = \tilde{A}\tilde{x} + \tilde{f}(\tilde{x}, \tilde{y}) \\ \dot{\tilde{y}} = \tilde{B}\tilde{y} + \tilde{g}(\tilde{x}, \tilde{y}) \end{cases}; \quad \tilde{A} = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{(c+p)\times(c+p)}; \quad \tilde{f} = \begin{pmatrix} f \\ 0 \end{pmatrix}. \tag{4.36}$$

Here, $\tilde{g} = g$ and $\tilde{B} = B$. Further, note that span $\left\{ \begin{pmatrix} x \\ 0 \end{pmatrix} \right\}$ is an invariant subspace for $\tilde{A}$. The eigenvalues of $\tilde{A}$ are the same as those of $A$ and an additional zero with multiplicity $p$, thus there are $c + p$ center directions and $d$ hyperbolic directions. Applying the center manifold theorem to the fixed point $0 \in \mathbb{R}^{c+p+d}$ of (4.36) we obtain that there exists a $\mathcal{C}^{r-1}$ manifold $\tilde{W}^C(0)$ tangent to $E^C$ at $\begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ which is invariant and is of dimension $c+p$. The geometry of this manifold is illustrated in Fig. 4.5. Note in the figure that there is no dynamics in the $\varepsilon$ direction, as $\dot{\varepsilon} = 0$ and that $(x, y) = (0, 0) \in \mathbb{R}^{c+d}$ remains a fixed point for $\varepsilon \neq 0$.

Computing $\tilde{W}^C(0)$ is done in a similar fashion as before. We use the center manifold theorem to get

$$\tilde{y} = y = \tilde{h}(\tilde{x}) = \tilde{h}(x, \varepsilon) = \mathcal{O}(\|x\|^2, \|\varepsilon\|\|x\|, \|\varepsilon\|^2) = \mathcal{O}(\|x\|^2, \|\varepsilon\|\|x\|). \tag{4.37}$$

The order $\|\varepsilon\|^2$ term was dropped as $x = 0$ must remain a fixed point. The function $h$ describes the graph of $W^C_\varepsilon(0)$. Then the reduced dynamics on $W^C_\varepsilon(0)$ is

$$\boxed{\dot{x} = Ax + f(x, \tilde{h}(x, \varepsilon), \varepsilon).} \tag{4.38}$$

This can then be applied to the perturbed example from above.

*Example* 4.4 (Revisting the perturbation). Recall the dynamical system

$$\begin{cases} \dot{x} = xy + \varepsilon x \\ \dot{y} = -y + \alpha x^2. \end{cases} \tag{4.39}$$

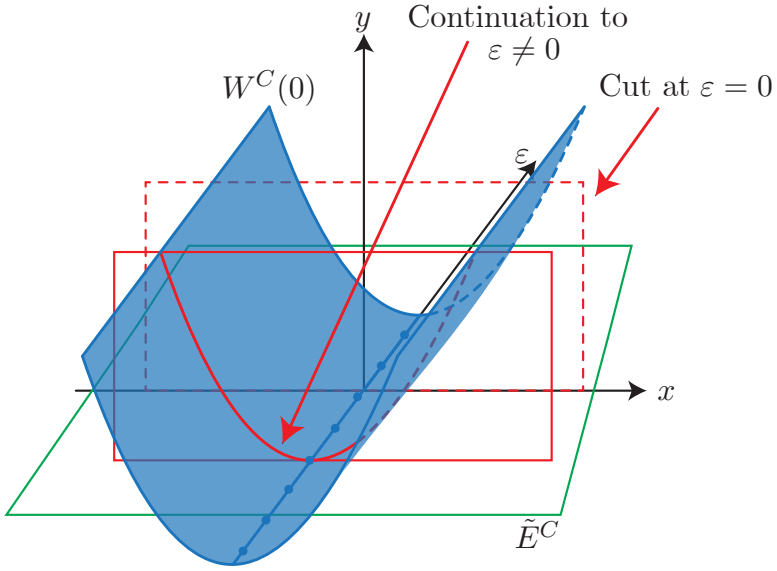


Figure 4.5: Geometry of the center manifold with the perturbation.

We have the persisting fixed point at $(x, y) = (0, 0) \in \mathbb{R}^{c+d}$ and the system is already in standard form with

$$A = 0; \quad B = -1; \quad f(x, y, \varepsilon) = xy + \varepsilon x; \quad g(x, y, \varepsilon) = \alpha x^2. \tag{4.40}$$

We apply the center manifold theorem and get the existence of $W^C_\varepsilon(0)$ for $|\varepsilon| \ll 1$. This manifold satisfies

$$y = \tilde{h}(x, \varepsilon) = ax^2 + bx\varepsilon + c\varepsilon^2 + dx^3 + ex^2\varepsilon + jx\varepsilon^2 + k\varepsilon^3 + lx^4 \ldots. \tag{4.41}$$

The term $c\varepsilon^2$ must be equal to 0 for all $\varepsilon$ such that the fixed point persists, therefore $c = 0$. Next the invariance $y(t) = \tilde{h}(x(t)), \varepsilon)$ is used, taking the time derivative on both sides yields

$$\dot{y} = [2ax + b\varepsilon + \mathcal{O}(2)] \underbrace{[\mathcal{O}(3) + \varepsilon x]}_{=\dot{x} \text{ from ODE and (4.41)}} = 2a\varepsilon x^2 + b\varepsilon^2 x + \mathcal{O}(4). \tag{4.42}$$

The $\mathcal{O}(n)$ designates terms of total degree $n$, for example $x^n$ or $x^{n-k}\varepsilon^k$. From the ODE we find

$$\dot{y} = -y + \alpha x^2 = -ax^2 - bx\varepsilon - c\varepsilon^2 - dx^3 - ex^2\varepsilon - jx\varepsilon^2 - k\varepsilon^3 - \mathcal{O}(4) + \alpha x^2. \tag{4.43}$$

Comparing equal powers in these two equations we find

$$\mathcal{O}(x^2): \ 0 = \alpha - a; \qquad \mathcal{O}(\varepsilon^2): \ 0 = -c; \qquad \mathcal{O}(x\varepsilon): \ 0 = -b; \tag{4.44a}$$

$$\mathcal{O}(\varepsilon^3): \ 0 = -t; \qquad \mathcal{O}(x^3): \ 0 = -d; \qquad \mathcal{O}(x^2\varepsilon): \ 2a = -e; \tag{4.44b}$$

$$\mathcal{O}(x\varepsilon^2): \ b = -j. \tag{4.44c}$$

Thus the shape of $W_\varepsilon^C(0)$ is given by

$$y = \tilde{h}(x, \varepsilon) = \alpha(1 - 2\varepsilon)x^2 + \mathcal{O}(4). \tag{4.45}$$

The dynamics on $W_\varepsilon^C(0)$ is

$$\dot{x} = \varepsilon x + \alpha(1 - 2\varepsilon)x^3 + \mathcal{O}(5). \tag{4.46}$$

We can see there is no substantial change in the shape of $W_\varepsilon^C(0)$ relative to the $\varepsilon = 0$ case. The stability type is determined by the sign of $\varepsilon$ and a two time-scale dynamic persists.

From this example we may still wonder what effect the higher order terms have on the center manifold.

## 4.4 Normal forms

For a general treatment see [Guckenheimer and Holmes, 1990], here we will consider one example to illustrate the idea originally coming from Poincaré.

*Example* 4.5 (Reduced dynamics on 1-dimensional manifold). Consider the following 1-dimensional dynamical system

$$\dot{x} = x(\mu - x^2) + x^4; \quad 0 \leq |\mu| \ll 1. \tag{4.47}$$

The fixed points are at $x = 0$ and the roots of $g_\mu(x) = \mu - x^2 + x^3$. This function $g_\mu$ is illustrated in Fig. 4.6. By plotting $x$ as a function of $\mu$ such that $g_\mu(x) = 0$ we get the *bifurcation diagram* as shown in Fig. 4.7.

The fold bifurcation (see caption of Fig. 4.7) is created by quartic (order 4) terms, which become more important away from the origin. The pitchfork bifurcation is already captured by the cubic truncation. We would like to know when the truncation captures the full dynamics correctly near the origin. Poincaré showed that, in fact, the truncated system is topologically equivalent to the full system near the origin by using a change of coordinates to remove $\mathcal{O}(4)$ terms.

(i) Let $x = y + h_4(y) = y + ay^4 + \mathcal{O}(y^5)$, which is near the identity near the origin, hence it is also invertible near the origin (by the Implicit Function Theorem). Further, this preserves the ODE up to the $\mathcal{O}(3)$ terms.

(ii) Plug in $x(t)$ and $y(t)$ and take the derivative with respect to time to get

$$\dot{x} = \dot{y}(1 + 4ay^3 + \mathcal{O}(y^4)). \tag{4.48}$$

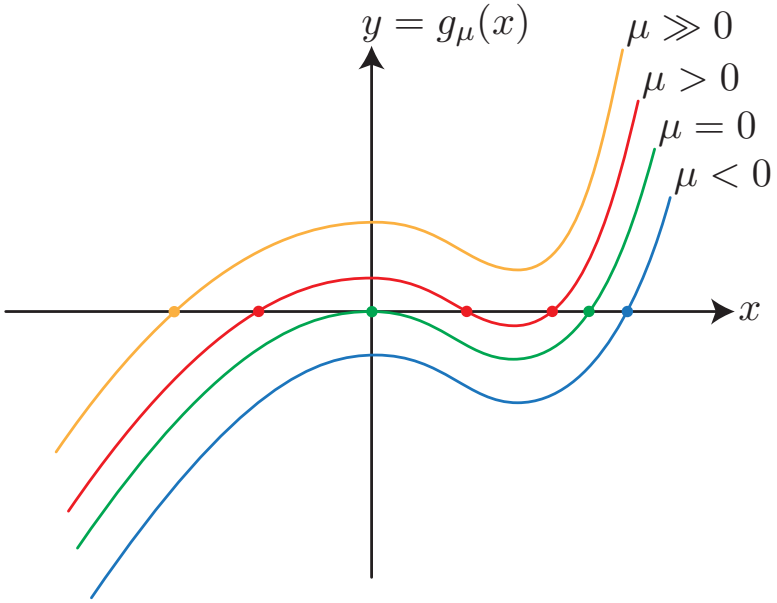


Figure 4.6: The functions $g_\mu$ for different values of $\mu$.

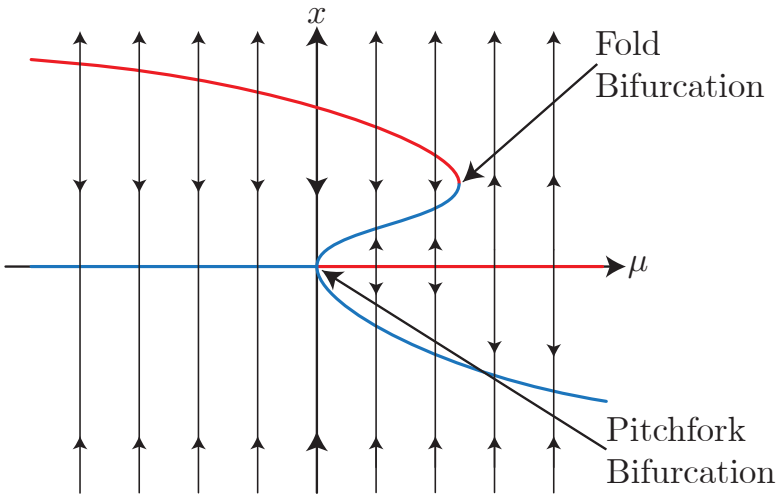


Figure 4.7: Bifurcation diagram for the 1-dimensional dynamical system. Red and blue demarcate if the fixed point at the given $(x, \mu)$ pair is stable (blue) or unstable (red). The curve is given by implicitly solving $g_\mu(x) = 0$.

(iii) Now use the ODE and find

$$\dot{x} = \mu x - x^3 + x^4 = \mu(y + ay^4) - (y + ay^4)^3 + (y - ay^4)^5 + \dots . \tag{4.49}$$

(iv) Combine the previous two steps and calculate

$$\dot{y} = \left[1 + 4ay^3 + \mathcal{O}(y^4)\right]^{-1} \left[\mu y + a\mu y^4 - y^3 + y^4 + \mathcal{O}(5)\right]. \tag{4.50}$$

At this point recall the alternating series (it could be verified with Taylor expansion)

$$\frac{1}{1+z} = 1 - z + \mathcal{O}(z^2); \quad 0 \leq |z| \ll 1. \tag{4.51}$$

We also note that a generalized result holds for an operator $A \in R^{m\times m}$, which could be used in higher dimensional calculations. This is referred to as the Neumann series. Applying this to the left term in the formula for $\dot{y}$ yields

$$\left[1 + 4ay^3 + \mathcal{O}(y^4)\right]^{-1} = 1 - 4ay^3 + \mathcal{O}(y^4). \tag{4.52}$$

Therefore we find

$$\dot{y} = \mu y - y^3 + y^4 \underbrace{(-4a\mu + a\mu + 1)}_{\text{choose } a \text{ such that this } =0} + \mathcal{O}(y^5). \tag{4.53}$$

We may now choose the parameter $a$ to make the coefficient of the quartic term 0. The $a$ that fulfills this is $a = \frac{1}{3\mu}$, using this we find

$$\boxed{\dot{y} = \mu y - y^3 + \mathcal{O}(y^5).} \tag{4.54}$$

This transformation has removed the quartic terms from the equation.

(v) Now we remove the $\mathcal{O}(y^5)$ terms similarly. First set

$$y = \xi + h_5(\xi) = \xi + b\xi^5 + \mathcal{O}(\xi^6) \tag{4.55}$$

and then continue as before, but with $y$ now playing the role of $x$ and $\xi$ playing the role of $y$.

(vi) The successive sequence of near identity coordinate changes turns out to converge usually. In general, it depends on the type of problem, sometimes resonant terms are not removeable and must stay as they are crucial to the dynamics. These resonant terms depend only on the linear part of the RHS, for more information see Chapter 3 of [Guckenheimer and Holmes, 1990].

Thus

$$\boxed{\dot{x} = \mu x - x^3} \tag{4.56}$$

is the *normal form* for the ODE for the study of bifurcations at the origin for $0 \leq \mu \ll 1$. It is topologically equivalent to the full system near $x = 0$ and captures the pitchfork bifurcation.

## 4.5 Bifurcations

A *bifurcation* is a qualitative change in the dynamical system

$$\boxed{\dot{x} = f(x, \mu); \quad x \in \mathbb{R}^n; \quad \mu \in \mathbb{R}^p.} \tag{4.57}$$

Linear stability analysis led to reducing to the center manifold (depending on parameters). From there we moved to normal forms which enable the analysis of qualitative dynamics under varying parameters.

**Definition 4.2** (Local bifurcation)**.** A *local bifurcation* is a qualitative change in the behavior of a system near a non-hyperbolic equilibrium. More precisely, a *bifurcation* occurs in (4.57) at $\mu = \mu_0$ near the fixed point $x = 0$ if there exists no neighborhood of $x = 0$ in which $\dot{x} = f(x, \mu_0)$ is topologically equivalent to all systems $\dot{x} = f(x, \mu)$ for $\|\mu - \mu_0\|$ small enough.

This idea of a bifurcation can be illustrated by a bifurcation surface which separates the space of dynamical systems into components. Within each component the dynamical systems are topologically equivalent, however elements from separate components are not. This is sketched in Fig. 4.8.

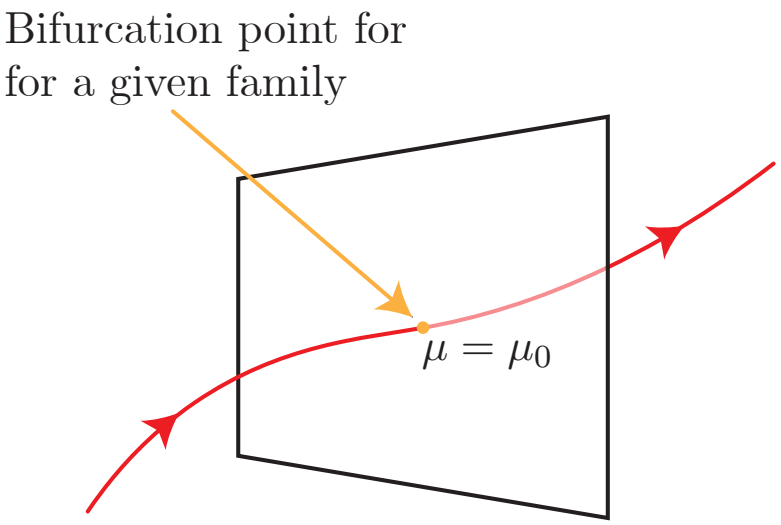


Figure 4.8: Illustration of a bifurcation surface. The near side of the surface is one component, and the far side the other. The red path is a family of dynamical systems $\{\dot{x} = f(x, \mu)\}_{\mu \in \mathbb{R}^p}$, going through the bifurcation point $\mu_0$.

We wish to understand what happens in the case that a given family of dynamical systems is nongeneric (atypical). For instance in the case that the family forms a tangency to the bifurcation surface, in which case a bifurcation does not take place, hence the family of dynamical systems is not general enough to capture all possible dynamics.

*Example* 4.6 (Nongeneric family of dynamical systems). In comparison to the family taken previously consider

$$\dot{x} = -a^2 x - x^3; \quad \mu = -a^2 \leq 0. \tag{4.58}$$

For every $\mu$ which we consider, there is only one fixed point $x = 0$ and it is stable for all values of $\mu$. However, we have unwittingly missed the full picture, as for $\mu > 0$ there are three fixed points, one of which are unstable. This situation is shown in Fig. 4.9.

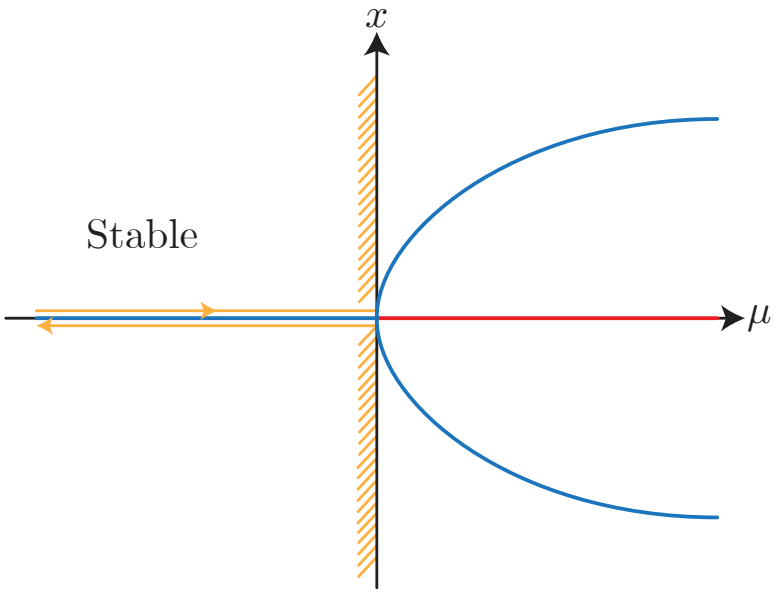


Figure 4.9: A nongeneric family of dynamical systems. We only see what happens to the left of the $x$-axis, and miss everything to the right, hence our family is tangent to the bifurcation surface.

This idea of nongeneric families motivates our next definition, which is also depicted in Fig. 4.10.

**Definition 4.3** (Universal unfolding)**.** A paramaterized family of dynamical systems crossing all nearby topological equivalence classes as the parameters vary is called a *universal unfolding*.

**Definition 4.4** (Codimension of a bifurcation)**.** The *codimension of a bifurcation* is the minimum number of parameters required for a universal unfolding. Thus a more degenerate bifurcation requires a larger codimension.

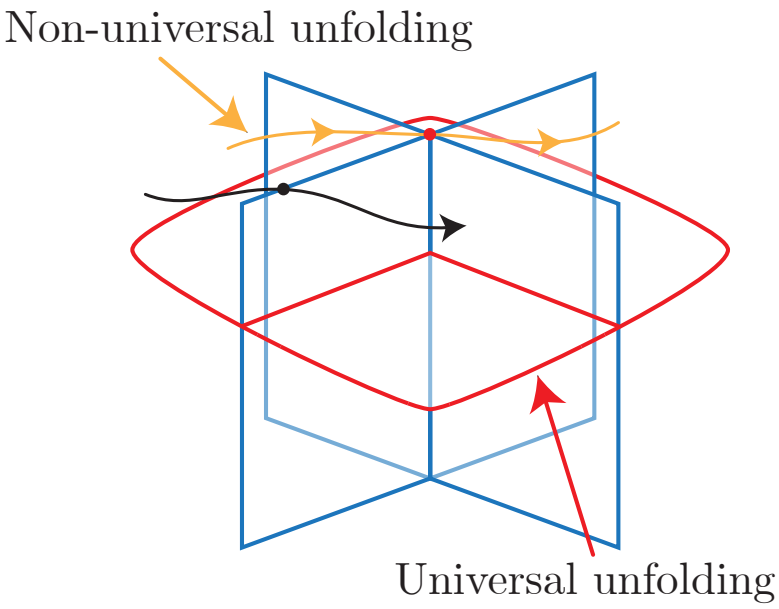


Figure 4.10: An example of universal unfolding (red) for the red bifurcation point which crosses the four topologically equivalent classes (components) created by the two bifurcation surfaces (blue). Furthermore, a nonuniversal unfolding is shown by the yellow 1-dimensional path at the top. Another universal unfolding in for the black bifurcation point, is shown by the 1-dimensional black family.

## 4.6 Codimension 1 bifurcations

We begin by considering different types of center manifolds. We start by classifying these by the number of zero eigenvalues which appear in the linearization.

(i) **Single zero eigenvalue** Our system is as follows

$$\dot{y} = f(y,\mu),\ y \in \mathbb{R}^n,\ \mu \in \mathbb{R}^p; \quad f(0,0) = 0; \quad \dim E^C = 1 \implies \dim W^C_\mu(0) = 1. \tag{4.59}$$

Thus we have a single zero eigenvalue, an example eigenvalue constellation which fulfills this is given in Fig. 4.11. The universal unfolding (generally parameterized by the normal form on $W^C_\mu(0)$) is given by

$$\boxed{\dot{x} = \tilde{\mu} \pm x^2; \quad x \in \mathbb{R}; \quad \tilde{\mu} \in \mathbb{R}.} \tag{4.60}$$

This is a codimension 1 bifurcation, i.e. only 1 parameter is needed for the universal unfolding. Hence we find a bifurcation diagram as shown in Fig. 4.12. This is called a *fold* or *saddle node* bifurcation, in this example it is *supercritical* as we show the $-x^2$

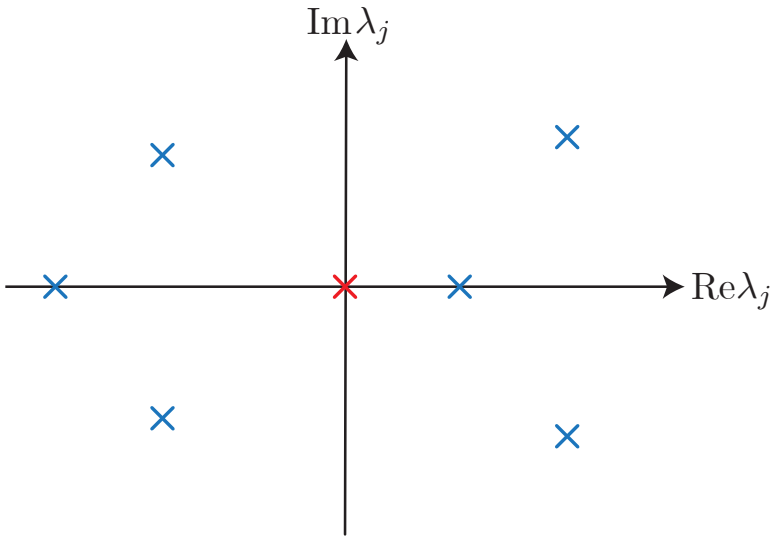


Figure 4.11: An example eigenvalue constellation with a single zero eigenvalue.

case. *Subcriticality* occurs when the diagram is mirrored with respect to the $x$-axis, i.e. when the fixed points disappear for $\mu > 0$. The sign which is actually used is problem dependent.

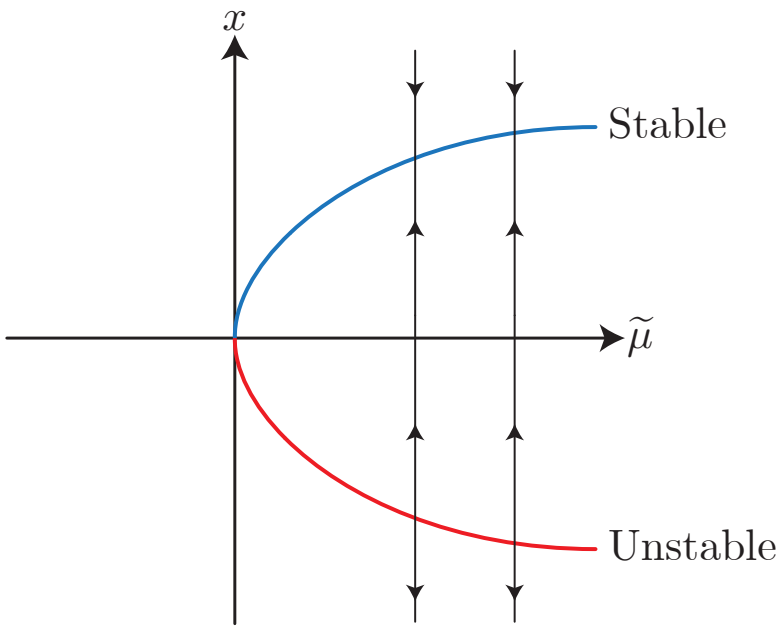


Figure 4.12: A codimension 1 fold bifurcation, which is supercritical.

*Remark* 4.5. *Hysteresis* is the dependence of the output on past input and is often related to interplay between two such fold bifurcations as shown in Fig. 4.13.

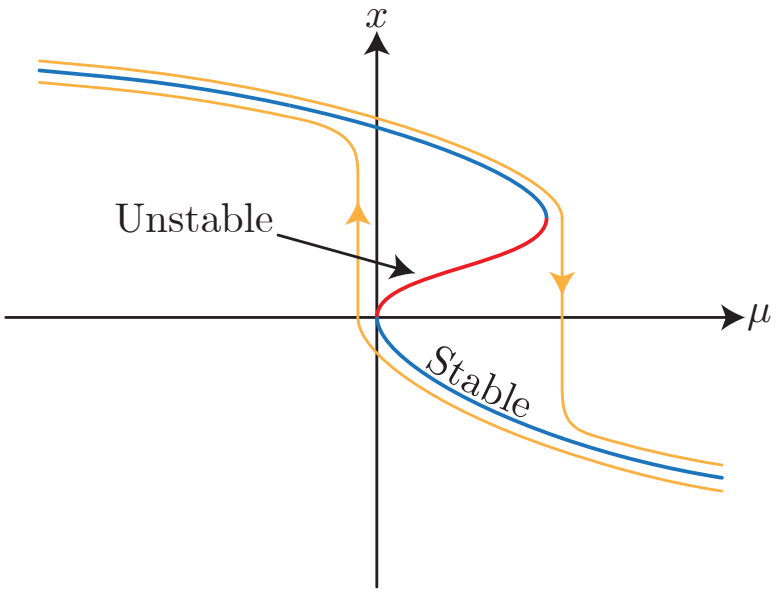


Figure 4.13: An example of hysteresis caused by two interacting fold bifurcations. In this case, the hysteresis is a result of purely deterministic dynamics, no additional memory (or delay) terms were needed.

(ii) **Single zero eigenvalue & origin remains a fixed point** Now we have more conditions and our setup is as follows

$$\dot{y} = f(y,\mu),\ y \in \mathbb{R}^n,\ \mu \in \mathbb{R}^p; \quad f(0,\mu) = 0,\ \forall \mu; \quad \dim E^C = 1 \implies \dim W^C_\mu(0) = 1. \tag{4.61}$$

Thus we have the additional degeneracy $D_\mu f(0,0) = 0$. We get the universal unfolding

$$\dot{x} = \tilde{\mu}x \pm x^2 = x(\tilde{\mu} \pm x); \quad x \in \mathbb{R}; \quad \tilde{\mu} \in \mathbb{R}. \tag{4.62}$$

This so called *transcritical bifurcation* is depicted in Fig. 4.14. In such bifurcations the observed stable state does not vary until the bifurcation, after which it begins to vary linearly.

*Example* 4.7 (A physical example of a transcritical bifurcation). The laser threshold is a physical example of a transcritical bifurcation [Strogatz, 2000, Haken, 1977]. The simplest model is of an active material held between two partially reflecting mirrors, with an energy

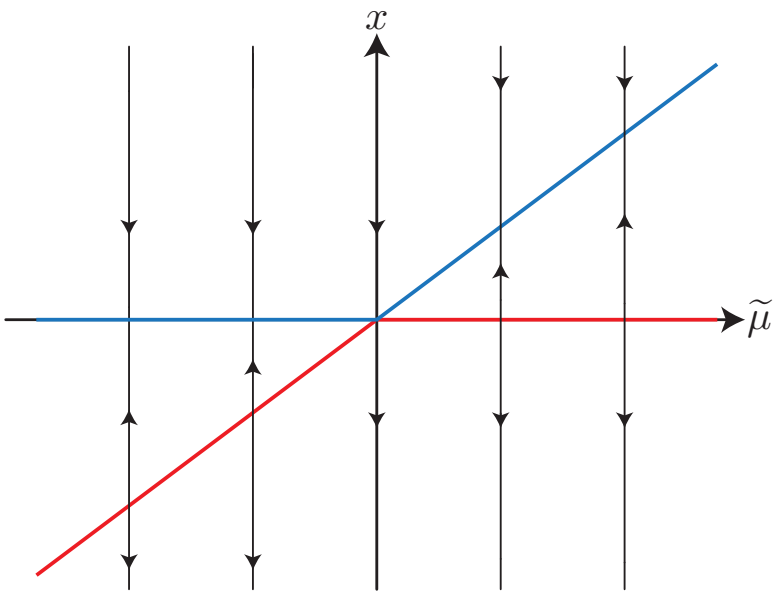


Figure 4.14: A transcritical bifurcation.

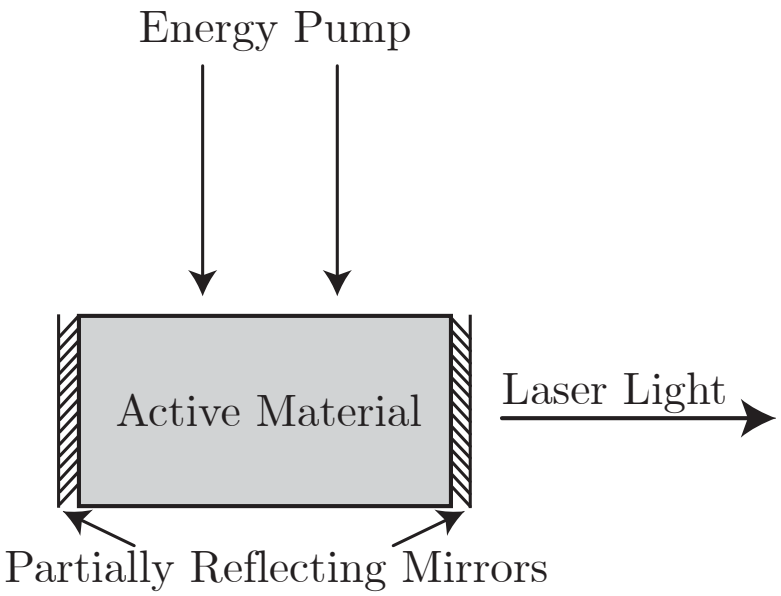


Figure 4.15: The simplest model of for the laser threshold example.

pump into the active material, from which a laser is emitted. This is depicted in Fig. 4.15.

The phenomenon we can observe in this system is that a cohert laser light is only gener-

ated over a certain pumping strength. Let $n(t)$ denote the number of photons emitted. Therefore $\dot{n}$ is equal to the gain minus the loss, i.e. if $N$ is the number of excited atoms and $G$ is the gain from stimulated emission, then

$$\dot{n} = GnN - kn. \tag{4.63}$$

The number of excited atoms $N$ decreases as an excited atom emits a photon. Therefore if $\alpha$ is the rate at which atoms drop back to the ground state and $N_0$ is the number of excited atoms (fixed without laser action)then we have

$$N(t) = N_0 - \alpha n. \tag{4.64}$$

Combining these two equations we find

$$\dot{n} = Gn(N_0 - \alpha n) - kn = \underbrace{(GN_0 - k)}_{=\mu} n - \alpha G n^2. \tag{4.65}$$

Hence, we have a transcritical bifurcation, and the laser threshold is at $\frac{k}{G}$, as shown in Fig. 4.16.

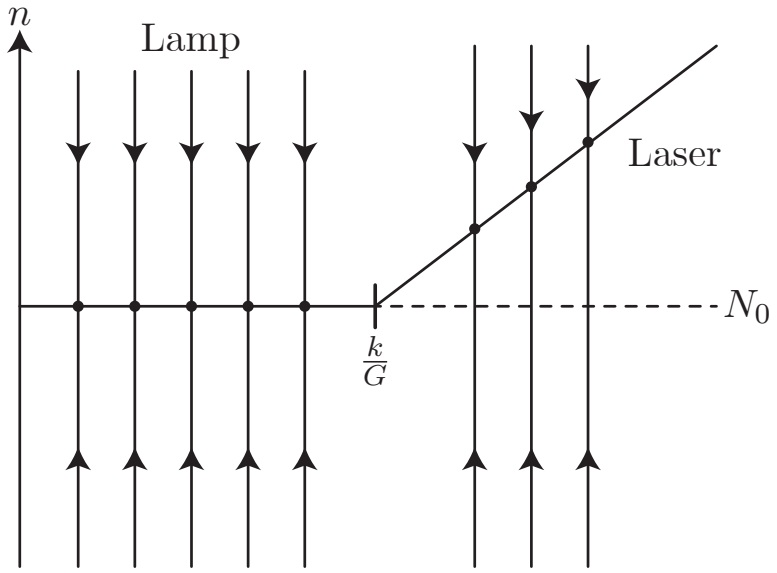


Figure 4.16: The laser threshold bifurcation at $\frac{k}{G}$.

(iii) **Single zero eigenvalue, origin remains a fixed point, & RHS is an odd function**
We have the same setup as previously, with the additional condition that

$$f(y,0) = -f(-y,0) \implies D_y^2 f(y,0) = -D_y^2 f(-y,0) \implies D_y^2 f(0,0) = 0. \tag{4.66}$$

This leads to the universal unfolding

$$\dot{x} = \tilde{\mu}x \pm x^3; \quad x \in \mathbb{R}; \quad \mu \in \mathbb{R}. \tag{4.67}$$

This type of bifurcation is called a *pitchfork bifurcation*. It is *supercritical* for the $-x^3$ case (cf. Fig. 4.17), and subcritical for the $+x^3$ case. The actual sign is problem dependent.

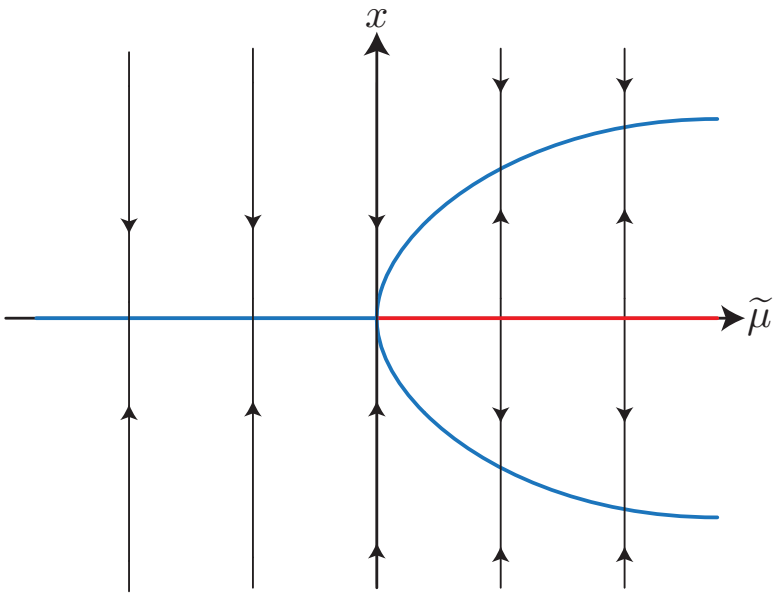


Figure 4.17: A supercritical (i.e with a $-$ sign) pitchfork bifurcation.

*Example* 4.8 (Buckling beam bifurcation). A beam fixed at the bottom with a downward force applied towards the top is an example of such a pitchfork bifurcation. As the force exceeds a critical threshhold, the middle position loses stability and two new symmetric, stable fixed points arise. This is illustrated in Fig. 4.18.

(iv) **Purely imaginary pair of eigenvalues** Now we have the following setup, with the eigenvalue constellation shown in Fig. 4.19.

$$f(0,0) = 0; \quad \dim E^C = 2; \quad \lambda_n(0) = \overline{\lambda}_{n-1}(0) = \pm i\omega \neq 0. \tag{4.68}$$

Such constellations are common in oscillatory systems at the loss of stability. This loss of stability is sketched in the right pane of Fig. 4.19. The reduced system dynamics on the 2-dimensional $W^C(0)$ for $\mu = 0$ is

$$\begin{pmatrix} \dot{u} \\ \dot{v} \end{pmatrix} = \begin{pmatrix} 0 & -\omega \\ \omega & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} \tilde{f}(u,v) \\ \tilde{g}(u,v) \end{pmatrix}. \tag{4.69}$$

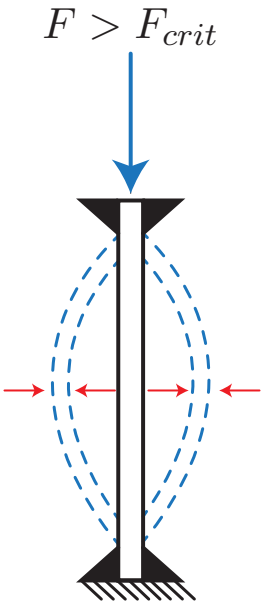


Figure 4.18: The buckling beam pitchfork bifurcation.

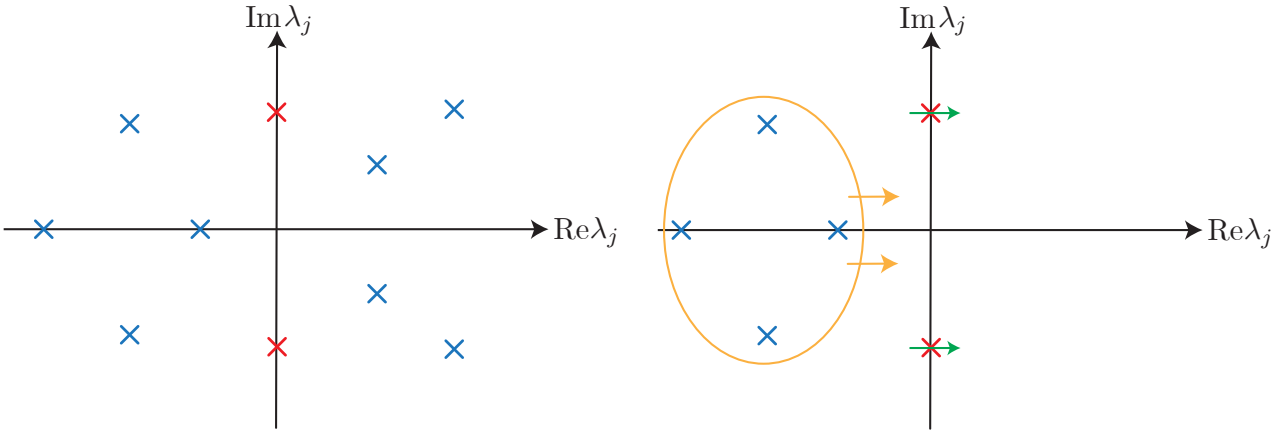


Figure 4.19: Left: Eigenvalue constellation for purely imaginary pair of eigenvalues. Right: Loss of stability in oscillating system as the pair of eigenvalues cross the imaginary axis.

The coordinates $(u, v)$ correspond to the basis given by $[\mathrm{Re}(e_n), \mathrm{Im}(e_n)]$. The normal form on $W^C_\mu(0)$ is

$$\dot{r} = r\left(d(\mu) + a(\mu)r^2\right) + \mathcal{O}(r^5); \qquad d(\mu) = \mathrm{Re}(\lambda_n(\mu)) \tag{4.70a}$$

$$\dot{\theta} = \omega(\mu) + e(\mu)r^2 + \mathcal{O}(r^4); \qquad \omega(\mu) = \mathrm{Im}(\lambda_n(\mu)). \tag{4.70b}$$

The functions $a$ and $e$ depend on the given $\tilde{f}$ and $\tilde{g}$ (thereby also on $f$). We then define

$$d_0 = \frac{d}{d\mu}\text{Re}\left[\lambda_n(\mu)\right]\big|_{\mu=0}\,; \quad \omega_0 = \text{Im}(\lambda_n(0)) \tag{4.71a}$$

$$a_0 = a(0) = \frac{1}{16}\left[\tilde{f}_{uuu} + \tilde{f}_{uvv} + \tilde{g}_{uuv} + \tilde{g}_{vvv}\right]\Big|_{(u,v)=0} \tag{4.71b}$$

$$+ \frac{1}{16\omega}\left[\tilde{f}_{uv}(\tilde{f}_{uu} + \tilde{f}_{vv}) - \tilde{g}_{uv}(\tilde{g}_{uu} + \tilde{g}_{vv}) - \tilde{f}_{uu}\tilde{g}_{uu} + \tilde{f}_{vv}\tilde{g}_{vv}\right]\Big|_{(u,v)=0}. \tag{4.71c}$$

Note that we must start from the standard form (4.69).

Building on this last case, we explore the extended center manifold in the next theorem.

**Theorem 4.6** (Hopf-Bogdanov)**.** *Assume the following*

*(i) The eigenvalues $\lambda_n$ and $\overline{\lambda}_n$ do cross the imaginary axis for $\mu \neq 0$, i.e. $d_0 \neq 0$.*

*(ii) The leading order nonlinearity in $\dot{r}$ is nonzero, i.e. $a_0 \neq 0$. Then there exists a unique extended center manifold $W^C_\mu(0)$ for $0 \leq |\mu| \ll 1$, on which the dynamical system is locally topologically equivalent to the universal unfolding*

$$\boxed{\dot{r} = r\left(d_0\mu + a_0 r^2\right)\,; \quad \dot{\theta} = \omega_0 + e_0 r^2 + b_0\mu,} \tag{4.72}$$

*for $b_0, e_0 \in \mathbb{R}$.*

From here we explore two cases for the signs of $d_0$ and $a_0$.

(i) First, when each are strictly positive $d_0, a_0 > 0$. This yields a subcritical pitchfork bifurcation in $r$ and a continued rotation in $\theta$. The bifurcation of $r$ and the dynamics of $u$ and $v$ for $\mu < 0$ are illustrated in Fig. 4.20, for $|\mu|$ and $|r|$ small enough. There is an increased sensitivity to perturbation as $\mu$ increases (the domain of attraction shrinks). This is shown in Fig. 4.21. Such behavior can be observed in the transition to turbulence in pipe flows. In that context, the equation is referred to as the Stuart-Landau equation.

(ii) Next, the two coefficients have opposite signs $d_0 > 0$ and $a_0 < 0$. This leads to the supercritical Hopf bifurcation and a stable limit cycle as shown in Fig. 4.22

There is a sudden development of stable periodic oscillation as $\mu$ is increased. Often the parameter $\mu$ is a steady speed in the system which is changed gradually. Examples of this are

- increasing wind speed leading to bridge oscillations and collapse,
- increasing flight speed leading to wing flutter,
- increasing driving speed leading to death wobble on motorcycles.

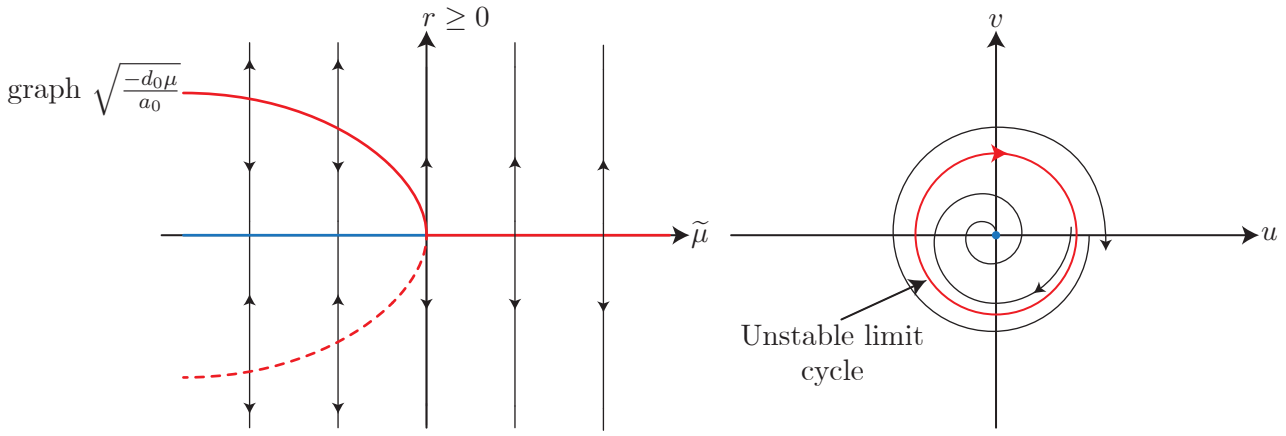


Figure 4.20: Left: The subcritical pitchfork bifurcation of $r$. The red arrow points towards the $r = \sqrt{\frac{-d_0\mu}{a_0}}$, the symmetric negative part of this curve is dashed, as negative values of $r$ are not relevant for this problem (the radius cannot be negative). Right: The effect of $\mu < 0$ on the system dynamics leading to an unstable limit cycle (red) in the $u - v$ plane.

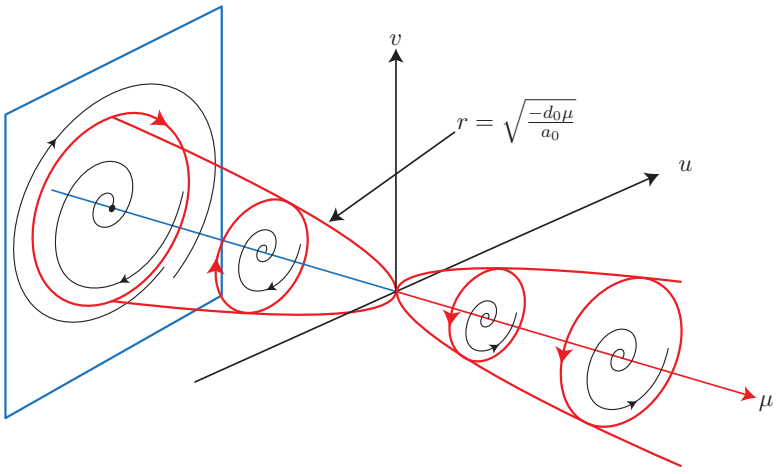


Figure 4.21: The full bifurcation diagram for the dynamical system in case (i).

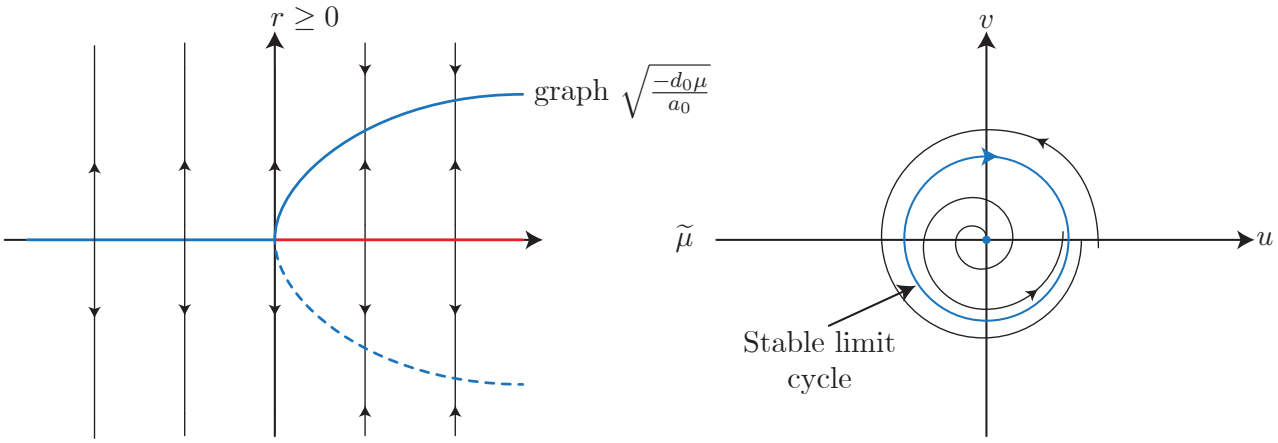


Figure 4.22: Left: The bifurcation in $r$, the blue arrow designates the curve given by $r = \sqrt{\frac{-d_0\mu}{a_0}}$, again the negative part is not relevant as the radius is always non-negative. Right: The stable limit cycle that arises for this system.

# Chapter 5

# Nonlinear dynamical systems on the plane

For this chapter the general setup will be

$$\dot{x} = f(x); \quad x \in \mathbb{R}^2; \quad f \in \mathcal{C}^1. \tag{5.1}$$

## 5.1 One degree of freedom conservative mechanical systems

Write $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, with $x_1$ denoting the position and $x_2 = \dot{x}_1$ denoting the speed. The total mechanical energy is given by $E(x) = \frac{1}{2}mx_2^2 + V(x_1)$, with the mass $m$ and the potential function $V$. Newton's law then gives the equations of motion

$$m\ddot{x}_1 = -\frac{dV(x_1)}{dx_1} \implies \begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = -\frac{1}{m}\frac{dV(x_1)}{dx_1}. \end{cases} \tag{5.2}$$

All of the forces derive from a time-independent potential. The energy is a first integral (conserved quantity of such systems), i.e.

$$\frac{d}{dt}E(x(t)) = \frac{\partial E}{\partial x_1}\dot{x}_1 + \frac{\partial E}{\partial x_2}\dot{x}_2 = 0. \tag{5.3}$$

Therefore on any given trajectory $x(t)$ the energy $E(x(t)) = E_0$ is constant. Hence we can derive an explicit equation for the velocity along a trajectory which suppresses the time dependence

$$x_2 = \pm\sqrt{\frac{2}{m}(E_0 - V(x_1))}. \tag{5.4}$$

This leads to multiple consequences for these systems

(i) Trajectories form symmetric pairs (w.r.t. the $x_1$-axis) of the same energy.

(ii) There is a clockwise orientation for trajectories due to

$$\dot{x}_1 = x_2 \implies \begin{cases} x_2 > 0 \implies x_1 \text{ increases} \\ x_2 < 0 \implies x_1 \text{ decreases.} \end{cases} \tag{5.5}$$

This is shown graphically in Fig. 5.1.

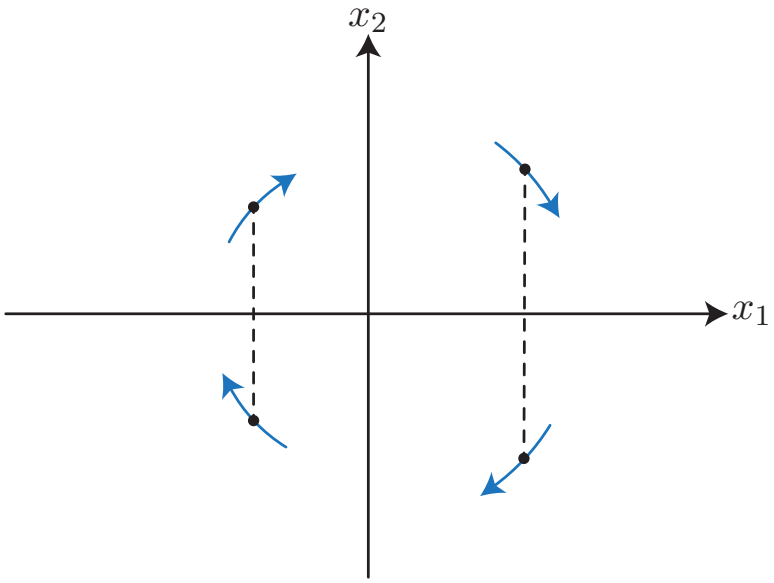


Figure 5.1: The clockwise orientation of trajectories in 1 degree of freedom conservative mechanical systems.

(iii) Local minima of the potential give rise to a center fixed point surrounded by closed orbits. The minimum $x^*$ is a fixed point if $x_2^* = 0$ (which occurs if $V(x_1^*) = E_0$) as $\dot{x}_2 \propto V'(x_1^*) = 0$. The trajectories of the closed orbits can be found using $x_2 = \pm\sqrt{\frac{2}{m}(E_0 - V(x_1))}$. Such a fixed point at a local minimum of the potential surrounded by closed orbits is shown in Fig. 5.2.

(iv) At local maxima of the potential saddle type fixed points arise. The maximum $x^*$ is a fixed point if $x_2^* = 0$ (which occurs if $V(x_1^*) = E_0$) as $V'(x_2^*) = 0$. The velocity can be

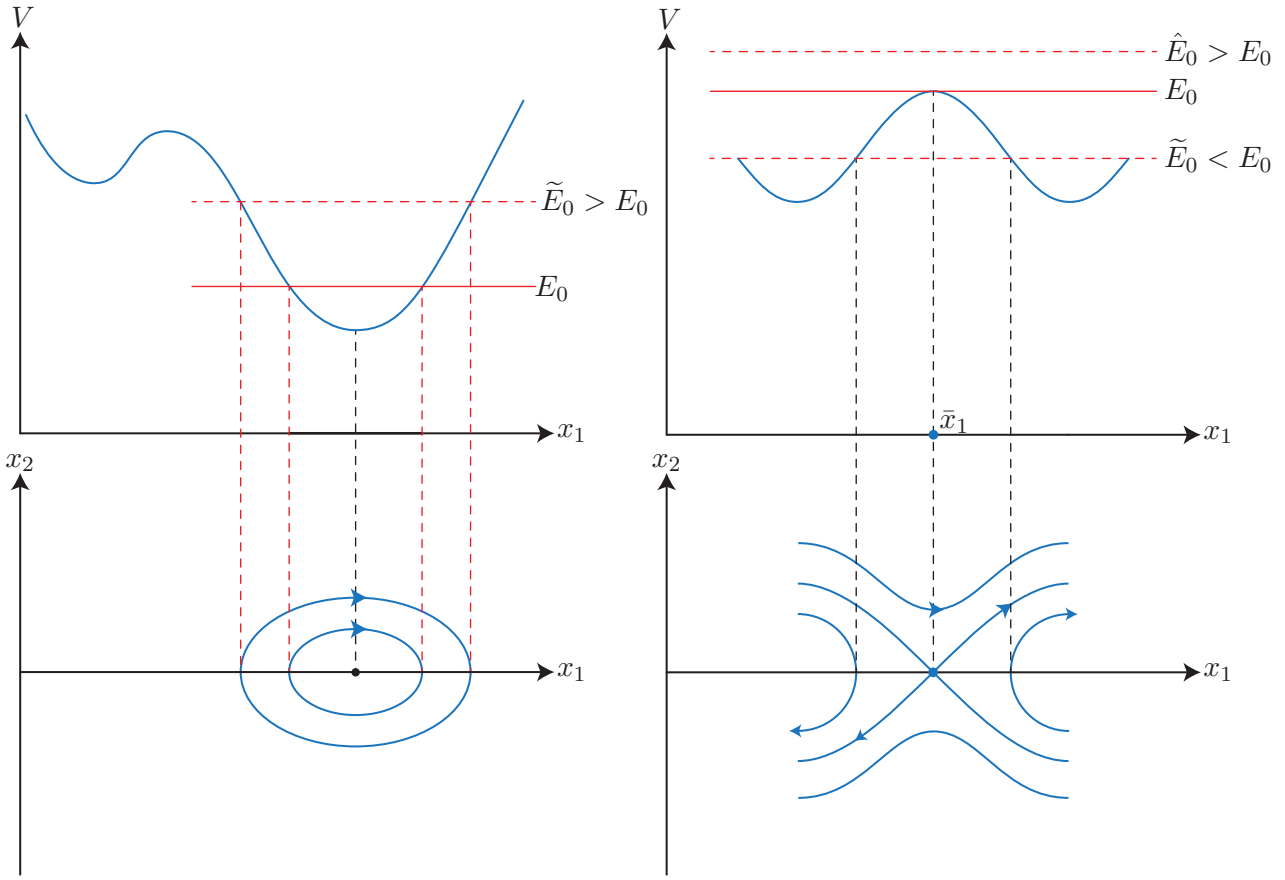


Figure 5.2: Left: Closed orbits around the fixed point at the local minimum of the potential function. Right: The saddle type fixed point at the local maxima of the potential. The clockwise orientations comes from consequence (ii).

approximated as follows

$$x_2 = \pm\sqrt{\frac{2}{m}(E_0 - V(x_1))} \tag{5.6a}$$

$$= \pm\sqrt{\frac{2}{m}}\sqrt{\underbrace{(E_0 - V(x_1^*))}_{=0} - \underbrace{V'(x_1^*)}_{=0}(x_1 - x_1^*) - \frac{1}{2}\underbrace{V''(x_1^*)}_{<0}(x_1 - x_1^*)^2 + \mathcal{O}((x_1 - x_1^*)^3)} \tag{5.6b}$$

$$= \pm\sqrt{\frac{2}{m}}\sqrt{-\frac{V''(x_1^*)}{2}}\,|x_1 - x_1^*| + \text{higher order terms.} \tag{5.6c}$$

The behavior is sketched in Fig. 5.2

(v) A local minimum has local maxima to the left and right, unless $V$ is monotonously nondecreasing to infinity. Then we have two cases to differentiate. One where the local maxima are at the same level, and the other where one local maximum has a smaller potential than the other (or in one direction $V$ monotonously increases towards infinity). *Homoclinic orbits* are special periodic orbits that are formed on a level set that contains a single saddle point. *Heteroclinic orbits* are formed when the level set contains two saddle points, i.e. when the two maxima have the same potential. These two cases are illustrated in Fig. 5.3.

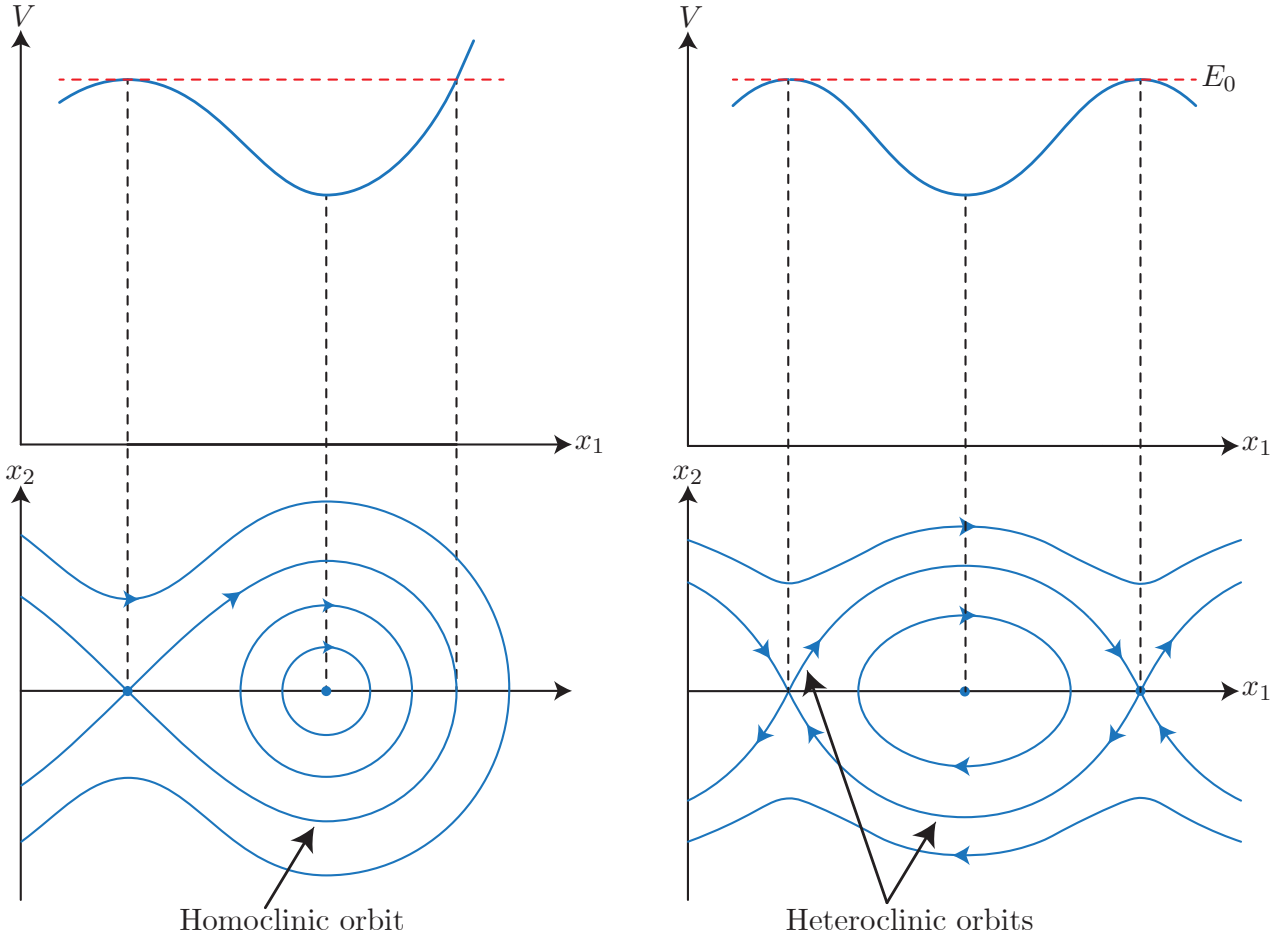


Figure 5.3: Left: One maximum has a lower potential than the other (or $V$ is nondecreasing towards infinity on the right). The black arrow denotes a homoclinic orbit. Right: All maxima have the same potential. The black arrows denote heteroclinic orbits.

These consequences are illuminated in a basic example.

*Example* 5.1 (The Duffing oscillator). Let the equation of motion be given by

$$\ddot{x} - x + x^3 = 0; \quad x_1 = x; \quad x_2 = \dot{x}. \tag{5.7}$$

Therefore we get the first order ODE

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = x_1 - x_1^3 = -V'(x_1) \end{cases} \implies E(x) = \frac{1}{2}x_2^2 + V(x_1) = \frac{1}{2}x_2^2 + \left(-\frac{1}{2}x_1^2 + \frac{1}{4}x_1^4\right). \tag{5.8}$$

The full phase portrait with two homoclinic orbits is shown in Fig. 5.4. We find two stable fixed points at the potential minima and they are separated by a potential maximum in the middle. Periodic orbits surround the stable fixed points. These orbits are separated from the rest of the phase space by the two homoclinic orbits of the saddle. Outside, we have periodic orbits that correspond to energy level-sets higher than the potential of the saddle.

## 5.2 Global behavior in two dimensional autonomous dynamical systems

Consider the following dynamical system

$$\dot{x} = f(x); \quad x \in \mathbb{R}^2; \quad f \in \mathcal{C}^1. \tag{5.9}$$

Further assume that solutions exist for all times, therefore there exists a unique solution $x(t; x_0)$ for all $t \in \mathbb{R}$ such that $x(0; x_0) = x_0$. Next, a few definitions are introduced enabling a deeper exploration of such systems.

**Definition 5.1.** A point $p \in \mathbb{R}^2$ is an *$\omega$-limit point* of $x_0$ if there exists a monotone increasing unbounded sequence of times $\{t_i\}_{i=1}^{\infty}$ with $t_1 \geq 0$ such that

$$\lim_{i \to \infty} x(t_i; x_0) = p. \tag{5.10}$$

**Definition 5.2.** A point $q \in \mathbb{R}^2$ is an *$\alpha$-limit point* of $x_0$ if it is an $\omega$-limit point in backward time.

**Definition 5.3.** The *$\omega$-limit set* of $x_0$, denoted $\omega(x_0)$ is the set of all $\omega$-limit points of $x_0$. The *$\alpha$-limit set* of $x_0$, denoted $\alpha(x_0)$ is the set of all $\alpha$-limit points of $x_0$.

*Remark* 5.1. Note that $\omega(x_0)$ and $\alpha(x_0)$ is the same for all $x_0$ along a given trajectory, thus limit sets can be associated with full trajectories.

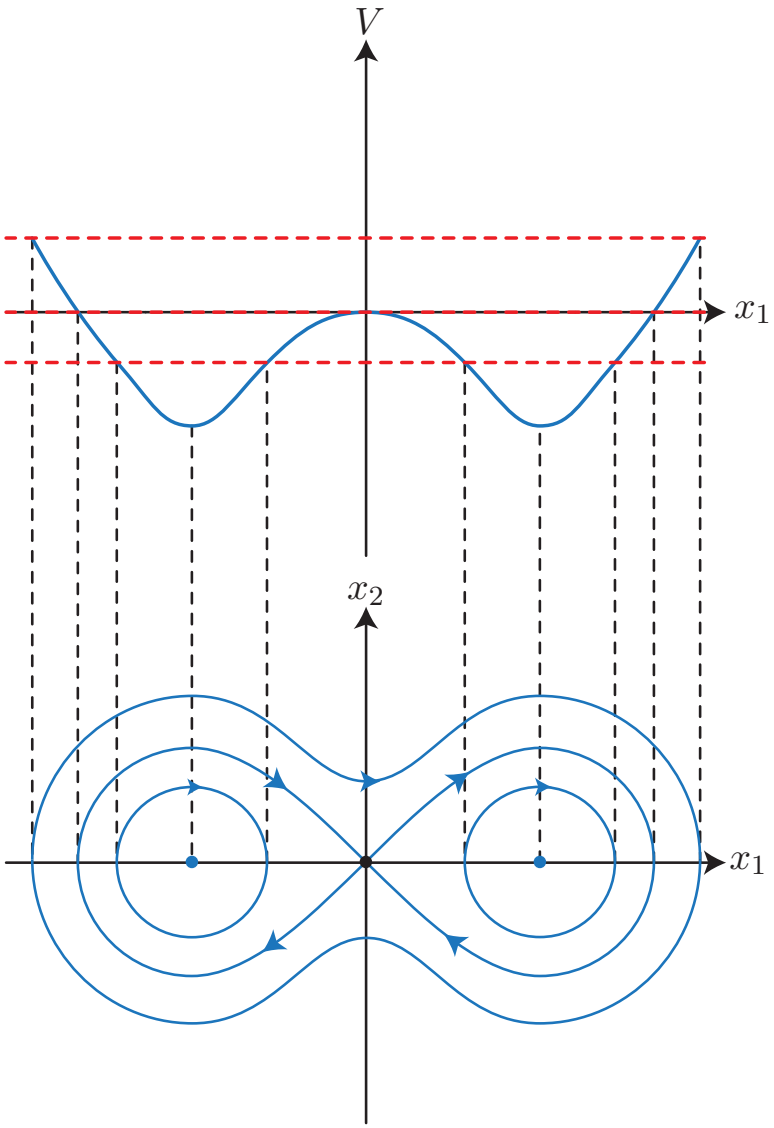


Figure 5.4: Phase portrait of the Duffing oscillator using the potential $V$.

*Example* 5.2 (Examples of limit sets). Below (Fig. 5.5) three different limit sets are depicted. For the leftmost dynamical system we have $\omega(p) = \alpha(p) = \{p\}$ and

$$\omega(x_0) = \emptyset, \qquad \omega(\tilde{x}_0) = \{p\}, \qquad \omega(\hat{x}_0) = \emptyset, \tag{5.11a}$$

$$\alpha(x_0) = \emptyset, \qquad \alpha(\tilde{x}_0) = \emptyset, \qquad \alpha(\hat{x}_0 = \{p\}. \tag{5.11b}$$

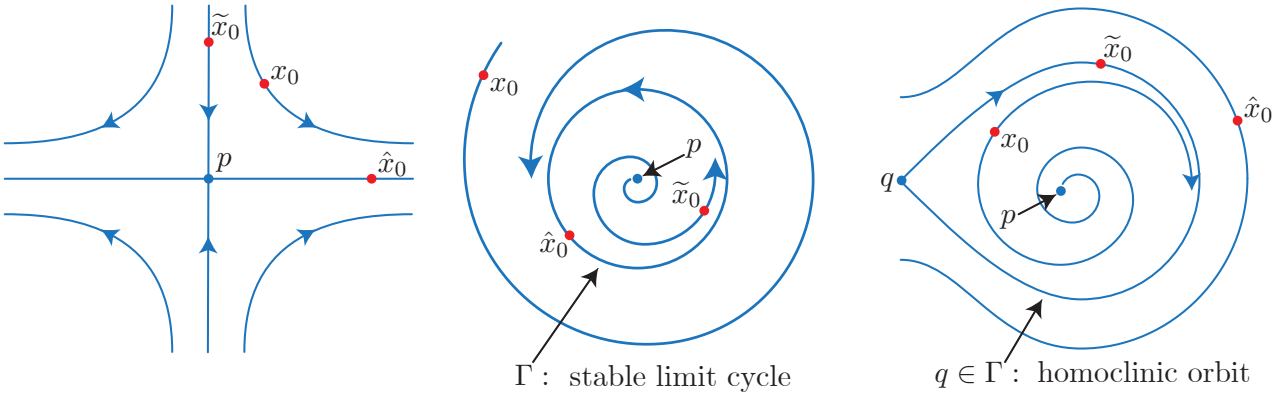


Figure 5.5: Three example systems for which we examine the respective limit sets. The arrows labelled $\Gamma$ denote a stable limit cycle (middle) and a homoclinic orbit (right).

For the middle dynamical system we have

$$\omega(x_0) = \Gamma, \qquad \omega(\tilde{x}_0) = \Gamma, \qquad \omega(\hat{x}_0) = \Gamma, \tag{5.12a}$$
$$\alpha(x_0) = \text{unclear}, \qquad \alpha(\tilde{x}_0) = \{p\}, \qquad \alpha(\hat{x}_0) = \Gamma. \tag{5.12b}$$

For the rightmost dynamical system we have

$$\omega(x_0) = \Gamma, \qquad \omega(\tilde{x}_0) = \{q\}, \qquad \omega(\hat{x}_0) = \text{unclear}, \tag{5.13a}$$
$$\alpha(x_0) = \{p\}, \qquad \alpha(\tilde{x}_0) = \{q\}, \qquad \alpha(\hat{x}_0) = \text{unclear}. \tag{5.13b}$$

We now introduce two theorems which give us more insight into these limit sets.

**Theorem 5.2.** *If $x(t; x_0)$ is bounded, then $\omega(x_0)$ and $\alpha(x_0)$ are nonempty, closed, and connected. Further, the omega and alpha limit sets are invariant sets, i.e. they consist of full trajectories.*

**Theorem 5.3** (Poincaré-Bendixon)**.** *If $x(t; x_0)$ is bounded, then $\omega(x_0)$ and $\alpha(x_0)$ must be one of the following*

*(i) A connected set of fixed points,*

*(ii) A limit cycle,*

*(iii) A set of fixed points and their homo-/heteroclinic orbits (illustrated in Fig. 5.6).*

*There is no complex (chaotic) liminting behavior in 2-dimensional autonomous nonlinear systems.*

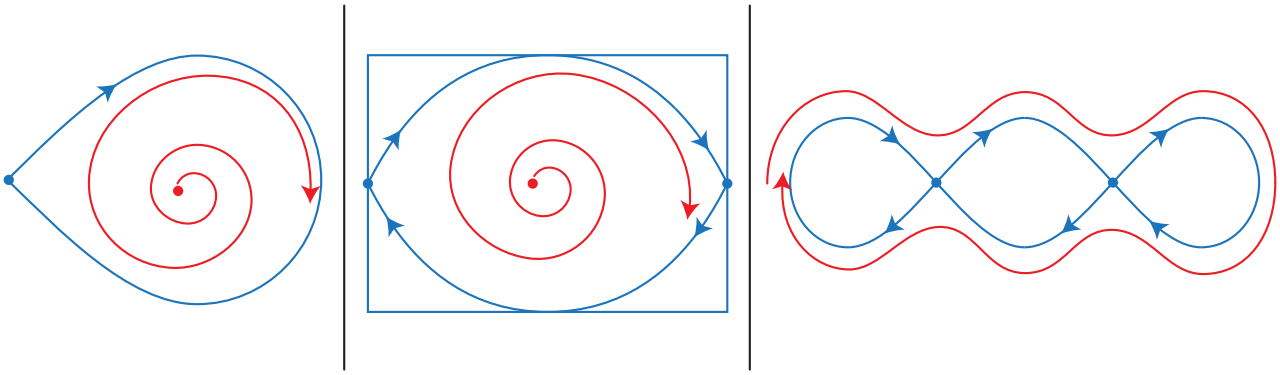

Figure 5.6: Examples of a set of fixed points connected by homo-/heteroclinic orbits.

*Remark* 5.4. The homo-/heteroclinic orbits are generally not robust, thus we should not expect to see them in a typical dynamical system. We can imagine that under small perturbations the continuous line of the orbit breaks up into separate unstable and stable manifolds as shown in Fig. 5.7. However, these orbits are robust within the class of conservative systems, and

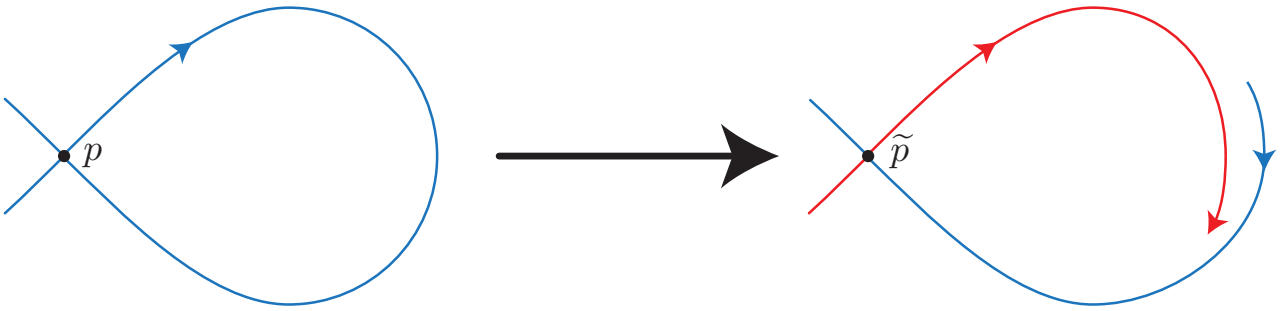


Figure 5.7: The loss of the homoclinic orbit under a small perturbation. Before the perturbation $\Gamma = W^S(p) = W^U(p)$ ($p$ is the hyperbolic fixed point), meanwhile after the perturbation $W^S(p) \neq W^U(p)$. The unstable manifold is given in red, and the stable manifold in blue.

conservative pertubations do not lead to global bifurcations as seen in Fig. 5.8.

*Remark* 5.5. A consequence of the Poincaré-Bendixon Theorem is that a forward-invariant, bounded open set without fixed points must contain a limit cycle. This is due to the elimination of possibilities (i) and (iii) in the theorem. An example of such a set is shown in Fig. 5.9.

**Proposition 5.6** (Bendixson criterion)**.** *Consider the following*

$$\dot{x} = f(x); \quad x \in \mathbb{R}^2; f \in \mathcal{C}^1; \quad U \text{ simply connected.} \tag{5.14}$$

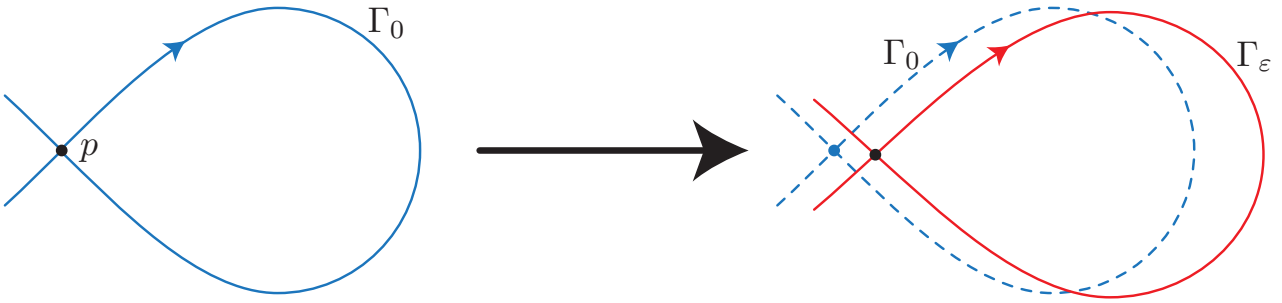


Figure 5.8: Small conservative perturbation does not lead to a global bifurcation in conservative systems. The energy level containing $\Gamma_0$ is robust, and perturbs to a new $\Gamma_\varepsilon$.

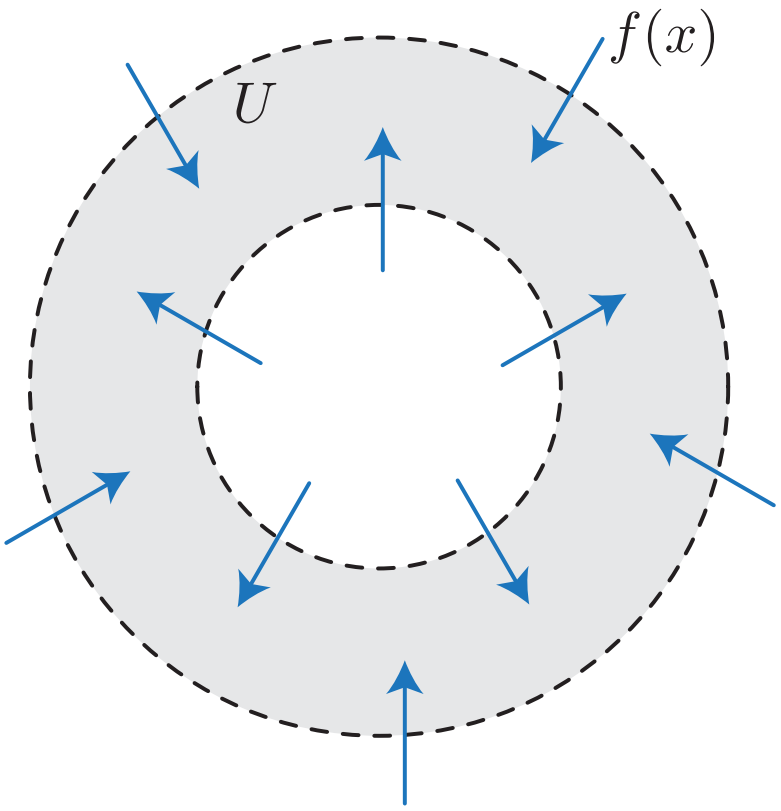


Figure 5.9: A forward-invariant, bounded open set without fixed points as in the remark. Such a set must contain a limit cycle.

*And further assume*

$$x \in U \implies \nabla \cdot f(x) \neq 0. \tag{5.15}$$

*This in turn implies $\nabla \cdot f(x)$ has constant sign on $U$. Then there cannot exist a limit cycle in*

$U$.

*Proof.* Assume, towards a contradiction, that there exists a closed orbit $\Gamma$ of period $\tau$. Calculate

$$f^{\perp}(x) = \begin{pmatrix} f_2(x) \\ -f_1(x) \end{pmatrix} \implies \oint_{\Gamma} f^{\perp} \cdot \underbrace{dr}_{=\dot{r}dt} = \int_0^{\tau} f^{\perp}(r(t)) \cdot f(r(t))dt = 0. \tag{5.16}$$

On the other hand we have by Green's Theorem

$$\oint_{\Gamma} f^{\perp} \cdot dr = \int\int_{\Gamma^{\circ}} \left(\nabla \times f^{\perp}\right)_z dA = -\int\int_{\Gamma^{\circ}} \underbrace{\nabla \cdot f}_{\neq 0} dA \neq 0. \tag{5.17}$$

Here, $\Gamma^{\circ}$ denotes the interior of the area enclosed by $\Gamma$. These two equations stand in contradiction to each other, therefore no limit cycle can exist in $U$. □

*Remark* 5.7. Another consequence of the Poincaré-Bendixson Theorem is that purely damped or purely forced perturbations of a conservative system cannot have limit cycles.

$$f(x) = \begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = -V'(x_1) + g(x_1, x_2) \end{cases}; \quad \nabla \cdot f = \nabla \cdot \begin{pmatrix} 0 \\ g(x) \end{pmatrix} = \frac{\partial g}{\partial x_2}. \tag{5.18}$$

If we assume pure damping or forcing, the divergence of $f$ cannot be 0 on any open set. If it is positive, we have pure forcing, alternatively if it is negative we have pure damping. In either case no limit cycle can exist.

# Chapter 6

# Time-dependent dynamical systems

Our setup for this chapter will be of nonautonomous $n$-dimensional systems, i.e. the following

$$\dot{x} = f(x,t); \quad x \in \mathbb{R}^n; \quad t \in \mathbb{R}. \tag{6.1}$$

Note that by adding a dimension to our phase space, we can make the system autonomous

$$X = \begin{pmatrix} x \\ t \end{pmatrix} \in \mathbb{R}^{n+1}; \quad F(X) = \begin{pmatrix} f \\ 1 \end{pmatrix} \implies \dot{X} = F(X). \tag{6.2}$$

## 6.1 Nonautonomous linear systems

For this section, consider

$$\dot{x} = A(t)x + b(t); \quad x \in \mathbb{R}^n; \quad A(t) \in \mathbb{R}^{n\times n}; \quad b(t) \in \mathbb{R}^n. \tag{6.3}$$

We separate (6.3) into two parts. The homogeneous part is

$$\dot{x} = A(t)x. \tag{6.4}$$

For the homogeneous part we have the solution $x(t) = \Phi_{t_0}^t x_0$, with the fundamental matrix solution $\Phi$ with the property $\Phi_{t_0}^{t_0} = I \in \mathbb{R}^{n\times n}$. The general solution of (6.3) is the sum of the general solution of (6.4) and a particular solution of (6.3). We can find the particular solution by, for example, the variation of constants formula. We refer to [Arnold, 1992] for the derivation. Thus we find the general solution to (6.3) to be

$$\boxed{x(t) = \Phi_{t_0}^t x_0 + \Phi_{t_0}^t \int_{t_0}^t \left(\Phi_{t_0}^s\right)^{-1} b(s)ds.} \tag{6.5}$$

Note that due to the group property of the flow map, the solution can be equivalently written as

$$x(t) = \Phi_{t_0}^t x_0 + \int_{t_0}^t \Phi_s^t b(s) ds. \tag{6.6}$$

From here we can see that it is enough to understand the homogeneous part. The stability of $x = 0$ in (6.4) is equivalent to the stability of $x_p(t) = \Phi_{t_0}^t \int_{t_0}^t \left(\Phi_{t_0}^s\right)^{-1} b(s) ds$ in (6.3). Unfortunately the general solution to (6.4) is unknown.

## 6.2 Time-periodic nonlinear systems

We now specify our system to be $T$-periodic, thus the evolution rule defined by $\dot{x} = A(t)x$ repeats itself with period $T$, i.e.

$$A(t+T) = A(t). \tag{6.7}$$

This implies

$$\Phi_{t_0}^{t_0+T} = \Phi_{t_0+T}^{t_0+2T} = \Phi_{t_0+2T}^{t_0+3T} = \ldots =: P_{t_0}, \tag{6.8}$$

The operator $P_{t_0}$ is called the *Poincaré map* (or time $T$ map) based at $t_0$. Thus we have

$$x_0(t_0 + nT) = \underbrace{P_{t_0} \ldots P_{t_0}}_{n \text{ times}} = P_{t_0}^n x_0. \tag{6.9}$$

The $n$-th iterate of the Poincaré map is denoted by $P_{t_0}^n$.

Time being a periodic variable allows us to study the geometry in the extended phase space $\mathbb{R}^n \times S^1$ as depicted in Fig. 6.1. Importantly, the stability of $x = 0$ can be studied as the stability of the $x = 0$ fixed point of the map $P_{t_0}$. Label the eigenvalues of $P_{t_0}$ $\rho_1, \ldots, \rho_n \in \mathbb{C}$ these are called *Floquet multipliers*. Since this is now a discrete dynamical system, we know how to assess the stability type of its fixed point. We have two cases

(i) If there exists an eigenvalue with $|\rho_i| > 1$ for $P_{t_0}$ or there exists a repeated eigenvalue with $|\rho_j| = 1$ for which the algebraic multiplicity is greater than the geometric multiplicity ($a > g$), then $x = 0$ is unstable for (6.4);

(ii) Otherwise, if each eigenvalue of $P_{t_0}$ has magnitude strictly less than 1, i.e. $|\rho_j| < 1$, then $x = 0$ is asymptotically stable for (6.4).

Floquet theory tells us that the following holds in general

$$\boxed{\Phi_{t_0}^t = B(t) e^{\Lambda(t-t_0)}; \quad B(t+T) = B(t) \in \mathbb{R}^{n\times n}; \quad \Lambda \in \mathbb{R}^{n\times n} \text{ is constant.}} \tag{6.10}$$

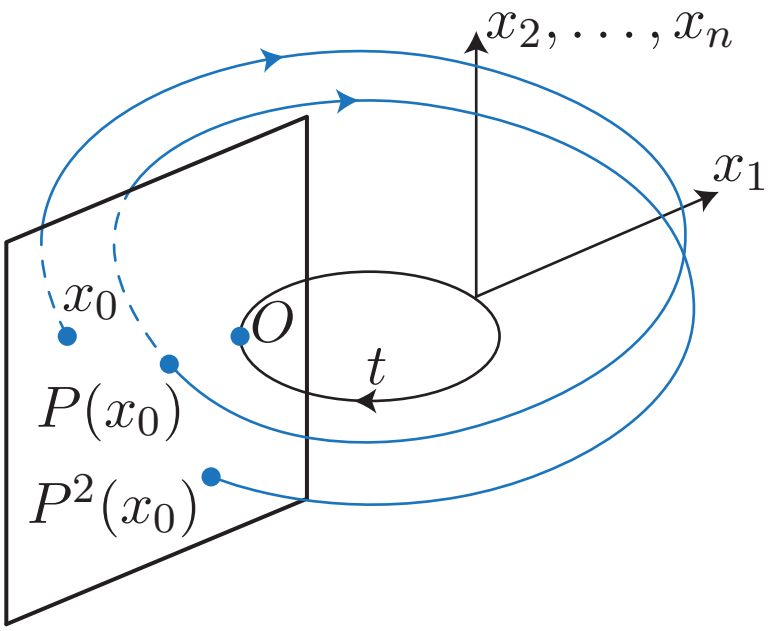


Figure 6.1: The map P records snapshots of the trajectory $x(t)$ taken at times $nT$. These snapshots are depicted along a single plane in the extended phase space. Due to periodicity, the extended phase space is made up as the product of the plane and the circle $S^1$. The snapshots shown are $Px_0 = x_1$ and $P^2x_0 = x_2$.

**Definition 6.1** (Floquet exponents)**.** The eigenvalues of $\Lambda$ are $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ and they are called *Floquet exponents*. The previously defined Floquet multipliers are related to $\lambda_j$ as

$$\boxed{\rho_j = e^{\lambda_j T} = e^{(\alpha_j + i\beta_j)T}.} \tag{6.11}$$

Hence, $\lambda_j$ is well-defined only up to addition of $i2k\pi/T$ for $k \in \mathbb{Z}$. Therefore we have

$$|\rho_j| \lessgtr 1 \Leftrightarrow \mathrm{Re}\lambda_j \lessgtr 0. \tag{6.12}$$

This is generally only numerically computable.

We now explore a sufficient criterion for instability of $x = 0$. For this, recall Liouville's Theorem (Abel's Theorem)

$$\boxed{\det(\Phi_{t_0}^t) = \det(\Phi_{t_0}^{t_0}) e^{\int_{t_0}^t \mathrm{Tr}[A(s)]ds}.} \tag{6.13}$$

This holds true for any dynamical system $\dot{x} = A(t)x$ with aperiodic time-dependence.

**Proposition 6.1** (Sufficient criterion for instability)**.** *If we have* $\int_0^T Tr[A(s)]ds > 0$ *then* $x = 0$ *is unstable for* $\dot{x} = A(t)x$ *with* $A(t) = A(t+T)$

*Proof.* Apply Liouville's Theorem to our $T$-periodic system to find

$$\prod_{j=1}^{n} \rho_j = e^{\sum_{j=1}^{n} \lambda_j T} = \det(P_{t_0}) = \det(\Phi_{t_0}^{t_0+T}) = e^{\int_{t_0}^{t_0+T} \mathrm{Tr}[A(s)]ds}. \tag{6.14}$$

Now we can compare the arguments of the exponential functions to derive

$$\boxed{\sum_{j=1}^{n} \lambda_j = \frac{1}{T} \int_0^T \mathrm{Tr}\left[A(s)\right] ds.} \tag{6.15}$$

Therefore our claim holds due to the case differentiation for stability above. Thus we can verify the stability without the numerical solution of (6.4). □

*Remark* 6.2. The change of variables $x = P(t)y$, for some matrix $P(t)$ transforms the system (6.4) into the form $\dot{y} = By$, an autonomous linear ODE.

Carrying out the change of coorinates, we obtain that

$$\dot{y} = P^{-1}(A(t)P - \dot{P})y. \tag{6.16}$$

But

$$\frac{d}{dt}\left(P(t)e^{Bt}\right) = APe^{Bt} \implies \dot{P} + PB = AP. \tag{6.17}$$

Together this implies $\dot{y} = By$.

## 6.3 Averaging

We will now attempt to study the mean evolution of oscillatory systems

$$\dot{x} = f(x,t); \quad x \in \mathbb{R}^n; \quad f(x,t+T) = f(x,t). \tag{6.18}$$

In general the averaged system

$$\dot{y} = \frac{1}{T} \int_0^T f(y,t)dt \tag{6.19}$$

does not capture the overall (mean) behavior of the system. This can be demonstrated in an example.

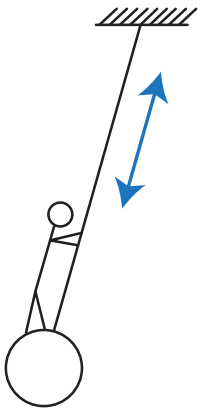

Figure 6.2: A swing (pendulum) with varying length and a mass $m$.

*Example* 6.1 (Averaging for a swing). Mechanically, we examine the stability of a pendulum with a periodically varying length as shown in Fig. 6.2. The first order model for this system is

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = -\omega_0^2(1 + a(t))x_1 \end{cases}; \quad a(t) = a(t+T), \quad 0 < |a| \ll 1. \tag{6.20}$$

Averaging these equations yields

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = -\omega_0^2(1 + \overline{a})x_1. \end{cases} \tag{6.21}$$

Where $\overline{a}$ is the average of $a(t)$, a constant. The phase portrait of this is known, as it is just a standard pendulum, i.e. $x = 0$ is stable. This means that the small forcing, when averaged, shouldn't move us from the stable position at $x = 0$. But we know that by periodically swinging, we don't just stay stuck at the equilibrium point of the swing, but instead are able to reach larger amplitude oscillations. Therefore, averaging clearly does not work in this situation.

However, there are situations in which averaging does work.

*Example* 6.2 (Averaging in slowly varying periodic systems). Slowly varying periodic systems are given by

$$\dot{x} = \varepsilon f(x, t, \varepsilon); \quad x \in \mathbb{R}^n; \quad 0 \leq \varepsilon \ll 1; \quad f(x, t, \varepsilon) = f(x, t+T, \varepsilon). \tag{6.22}$$

Such systems are also called *adiabatic* systems. The averaged system in this case is

$$\dot{y} = \varepsilon \frac{1}{T} \int_0^T f(y, t, 0)dt = \varepsilon \overline{f_0}(y). \tag{6.23}$$

Here only the leading order terms with respect to $\varepsilon$ in the integrand were taken, as $\varepsilon \ll 0$. Now we transform our original equation into the averaged equation via a near-identity, periodic change of variables

$$x = y + \varepsilon w(y,t); \quad w(y,t) = w(y,t+T). \tag{6.24}$$

Plugging this transformation into our differential equation, we find

$$\dot{x} = [I + \varepsilon D_y w]\,\dot{y} + \varepsilon \frac{\partial w}{\partial t} \tag{6.25a}$$

$$\dot{y} = \underbrace{[I + \varepsilon D_y w]^{-1}}_{I - \varepsilon D_y w + \mathcal{O}(\varepsilon^2)} \left[ \varepsilon \underbrace{f(y + \varepsilon w(y,t), t, \varepsilon)}_{=\dot{x}} - \varepsilon \frac{\partial w}{\partial t} \right] \tag{6.25b}$$

$$= \varepsilon \left[ f(y,t,0) - \frac{\partial w}{\partial t} \right] + \mathcal{O}(\varepsilon)^2. \tag{6.25c}$$

In the last equation we used the Taylor expansion. We cannot set $w(y,t) = \int_0^t f(y,s,0)ds$ as that would make $\omega$ aperiodic. Instead we separate $f$ into two parts, the average and the deviation, $f(y,t,0) = \overline{f}_0(y) + \tilde{f}(y,t)$. Note that the deviation has mean 0, therefore we set $w(y,t) = \int_0^t \tilde{f}(y,s)ds$ which is periodic. Therefore we find $\dot{y} = \varepsilon \overline{f}_0(y) + \mathcal{O}(\varepsilon^2)$, noting here that the order 2 term is still $T$ periodic.

This has given us that the original system is a $T$-periodic, $\mathcal{O}(\varepsilon^2)$ perturbation of the averaged system in the $y$ coordinates.

**Averaging Principle** Solutions of the original equation starting $\mathcal{O}(\varepsilon)$ close to those of the averaged system, stay $\mathcal{O}(\varepsilon)$ close for times of $\mathcal{O}\left(\frac{1}{\varepsilon}\right)$ (very long).

The averaging principle is depicted in Fig. 6.3.

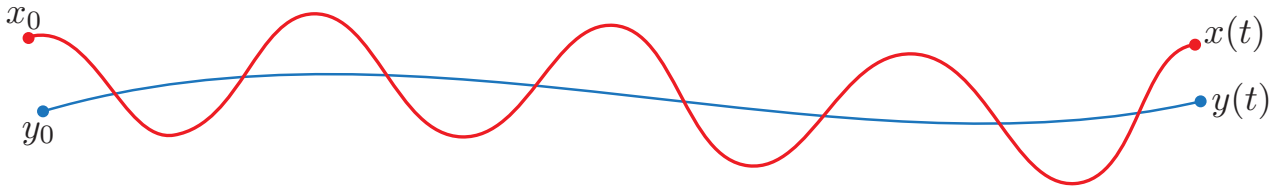


Figure 6.3: For the starting point $x_0$ which is $\mathcal{O}(\varepsilon)$ close to $y_0$, the trajectories remain order $\varepsilon$ close, i.e. $x(t)$ is $\mathcal{O}(\varepsilon)$ close to $y(t)$, for $t \ll \frac{K}{\varepsilon}$.

This has implications for the Poincaré map of the original equation

$$P^{\varepsilon}_{t_0} : x_0 \mapsto x(t_0+T, x_0; \epsilon); \quad P^{\varepsilon}_{t_0} = P^0_{t_0} + \mathcal{O}(\epsilon) \implies \left(P^{\varepsilon}_{t_0}\right)^n = \left(P^0_{t_0}\right)^n + \mathcal{O}(\varepsilon). \tag{6.26}$$

This only holds if $n \leq \frac{1}{\varepsilon}$. The time $T$ map of the averaged system was denoted as $P^0_{t_0}$.

*Example* 6.3 (Averaged system with a saddle-type fixed point ). Consider a system whose Poincaré map $P_0$ has a hyperbolic fixed point $p_0$. By the persistence of hyperbolic fixed points, the perturbed Poincaré map $P_\varepsilon$ also has a hyperbolic fixed point $p_\varepsilon$ which is $\mathcal{O}(\varepsilon)$ close to $p_0$. Furthermore, the $\mathcal{C}^1$ manifolds $W^U_{\text{loc}}(p_\varepsilon)$ and $W^S_{\text{loc}}(p_\varepsilon)$ are $\mathcal{O}(\varepsilon)$ close to their unperturbed counterparts. The unperturbed system and geometry of the perturbed system are shown in Fig. 6.4.

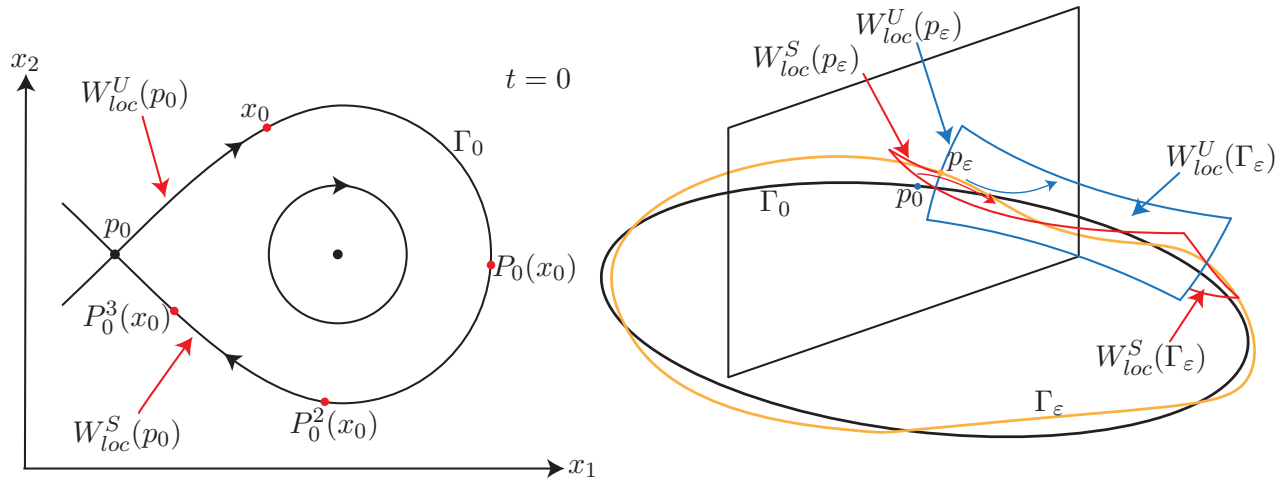


Figure 6.4: Left: The averaged system with the saddle type fixed point $p_0$. The upper arrow designates the local unstable manifold, and the lower arrow the local stable manifold. Right: The same in the extended phase space, time is represented by $S_1$ and runs clockwise. The hyperbolic limit cycle $\Gamma_\varepsilon$ is given by the yellow trajectory. Furthermore, the non-averaged local stable and unstable manifolds are drawn in red and blue respectively.

*Remark* 6.3. We can show that fixed points persist for $P_\varepsilon$ as follows. First we take the flow map for the system $F^t_{t_0}(y_0;\varepsilon)$, then we use the Taylor expansion to find

$$F^{t_0+T}_{t_0}(y_0;\varepsilon) = F^{t_0+T}_{t_0}(y_0;0) + \varepsilon \frac{\partial}{\partial \varepsilon} F^{t_0+T}_{t_0}(y_0;0) + \mathcal{O}(\varepsilon^2) \tag{6.27a}$$

$$= y_0 + \varepsilon \left.\frac{\partial y(t_0+T;t_0;\varepsilon)}{\partial \varepsilon}\right|_{\varepsilon=0} + \mathcal{O}(\varepsilon^2) \tag{6.27b}$$

$$= y_0 + \varepsilon \int_{t_0}^{t_0+T} \overline{f}(y_0)dt + \mathcal{O}(\varepsilon^2). \tag{6.27c}$$

In the last equality we used that

$$\left.\frac{\partial y(t_0+T;t_0,\varepsilon)}{\partial \varepsilon}\right|_{\varepsilon=0} = \left.\frac{\partial}{\partial \varepsilon}\int_{t_0}^{t_0+T} \dot{y}(s;t_0,\varepsilon)ds\right|_{\varepsilon=0} \tag{6.28a}$$

$$= \int_{t_0}^{t_0+T} \left.\frac{\partial}{\partial \varepsilon}\left(\varepsilon \overline{f}(y(s)) + \mathcal{O}(\varepsilon^2)\right)\right|_{\varepsilon=0} ds \tag{6.28b}$$

$$= \int_{t_0}^{t_0+T} \overline{f}(y_0)ds. \tag{6.28c}$$

From this we have

$$P_\varepsilon(y_0) - y_0 \Leftrightarrow y_0 + \varepsilon T\overline{f}(y_0) + \mathcal{O}(\varepsilon^2) - y_0 = 0 \Leftrightarrow T\overline{f}(y_0) + \mathcal{O}(\varepsilon) = 0. \tag{6.29}$$

And now we can use the Implicit Function Theorem to conclude.

*Example* 6.4 (Averaged system with a stable focus). Consider the averaged system with a stable focus around a fixed point $p_0$. The transition to the perturbed system is shown in Fig. 6.5.

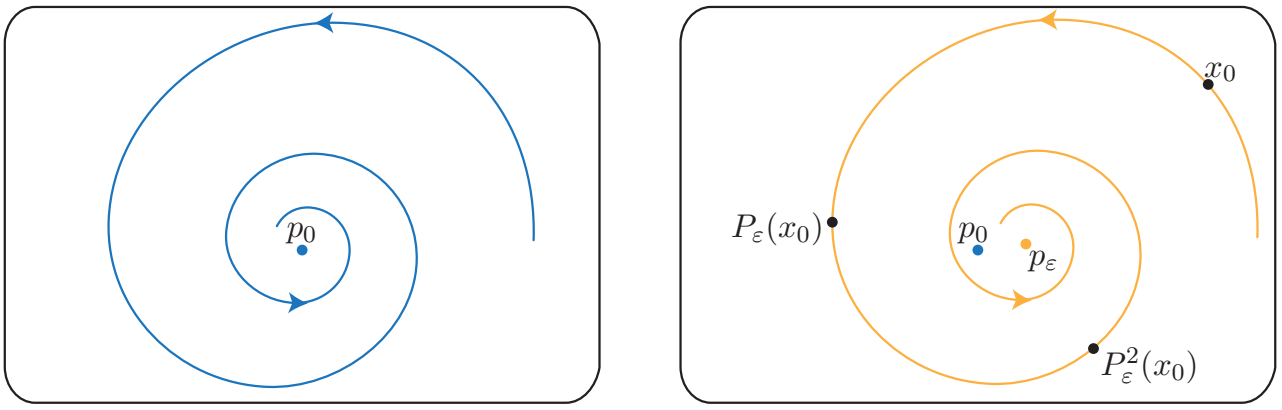


Figure 6.5: The stable focus around $p_0$ in the averaged system is preserved in the non-averaged system.

*Example* 6.5 (Averaged system with a limit cycle). Consider the averaged system with a limit cycle $\Gamma_0$. The non-averaged system also has a stable limit cycle $\Gamma_\varepsilon$ with dim $W^S(\Gamma_\varepsilon) = \mathbb{R}^{n+1}$ in extended phase space. Thus in the full system we have an attracting 2-dimensional invariant torus. The averaged and non-averaged system, along with this torus are shown in Fig. 6.6.

*Example* 6.6 (Weakly nonlinear oscillations). An important application of averaging is for weakly nonlinear oscillations. This is covered in-depth in [Guckenheimer and Holmes, 1990]. The system is given by

$$\ddot{x} + \omega_0^2 x = \varepsilon f(x,\dot{x},t) = \varepsilon f(x,\dot{x},t+T); \quad T = 2\pi\omega. \tag{6.30}$$

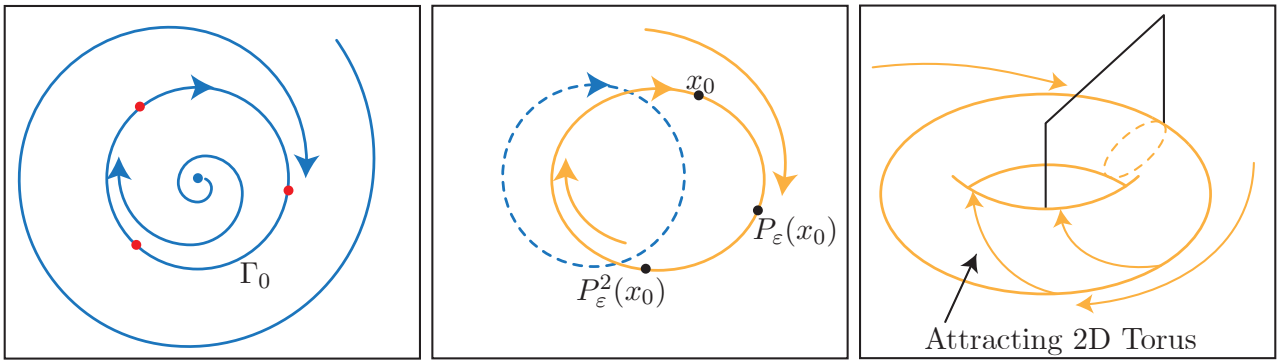


Figure 6.6: Left: The averaged system with the stable limit cycle $\Gamma_0$. Middle: The non-averaged system with the limit cycle $\Gamma_\varepsilon$. Right: The attracting 2-dimensional invariant torus (black arrow), trajectories can be seen as the thin yellow lines.

Averaging applies after a change of coordinates moving with the solutions of the $\varepsilon = 0$ limit. First transform the system to be a first-order ODE with $X = \begin{pmatrix} x \\ \dot{x} \end{pmatrix}$.

$$\dot{X} = AX + \varepsilon F(X,t); \quad A = \begin{pmatrix} 0 & 1 \\ -\omega_0^2 & 0 \end{pmatrix}; \quad F(X,t) = \begin{pmatrix} 0 \\ f(X_1, X_2, t) \end{pmatrix}. \tag{6.31}$$

We cannot directly apply averaging to (6.31). Instead we introduce a coordinate change along the trajectories of the system. Let $\Phi(t;\omega_0)$ be the fundamental matrix of solutions. In fact, this is explicitly computable for the harmic oscillator as

$$\Phi(t;\omega_0) = \begin{pmatrix} \cos(\omega_0 t) & -\sin(\omega_0 t) \\ -\omega_0 \sin(\omega_0 t) & -\omega_0 \cos(\omega_0 t) \end{pmatrix}. \tag{6.32}$$

It can be checked that $\dot{\Phi} = A\Phi$. Thus we change coordinates with $X(t) = \Phi(t,\omega_0)Y(t)$. As $\Phi$ is invertible, we have $Y(t) = \Phi^{-1}(t,\omega_0)X(t)$ with the inverse

$$\Phi^{-1}(t,\omega_0) = \begin{pmatrix} \cos(\omega_0 t) & -\frac{1}{\omega_0}\sin(\omega_0 t) \\ -\sin(\omega_0 t) & -\frac{1}{\omega_0}\cos(\omega_0 t) \end{pmatrix}. \tag{6.33}$$

Now we substitute this transformation into the differential equation (6.31) and use the fact that $\dot{\Phi} = A\Phi$ to find

$$\dot{\Phi}Y + \Phi\dot{Y} = A\Phi Y + \varepsilon F(\Phi^{-1}Y, t) \implies \dot{Y} = \varepsilon\Phi^{-1}F(\Phi Y, t). \tag{6.34}$$

This appears to be in the required form for averaging, except the right hand side is not generally periodic, because $\Phi$ has frequency $\omega_0$, while the driving has frequency $\omega$. It is only

periodic if $\omega = k\omega_0$ for $k \in \mathbb{Z}$, i.e. resonance of order $k$. Often it is important to analyze near-resonant oscillations, i.e. $\omega \approx k\omega_0$. Thus our expectation is to find an almost sinusoidal response of frequency $\frac{\omega}{k}$. Therefore we instead use the transformation $\Phi(t; \frac{\omega}{k})$ and find

$$\dot{Y} = \Phi^{-1}\left[A\Phi - \dot{\Phi}\right] Y + \varepsilon\Phi^{-1}F(\Phi Y, t). \tag{6.35}$$

Now write $Y = \begin{pmatrix} u \\ v \end{pmatrix}$, and we find

$$\begin{cases} \dot{u} = -\dfrac{k}{\omega}\left[\left(\dfrac{\omega^2 - k^2\omega_0^2}{k^2}\right) x + \varepsilon f(x, \dot{x}, t)\right] \sin\left(\dfrac{\omega t}{k}\right) \\ \dot{v} = -\dfrac{k}{\omega}\left[\left(\dfrac{\omega^2 - k^2\omega_0^2}{k^2}\right) x + \varepsilon f(x, \dot{x}, t)\right] \cos\left(\dfrac{\omega t}{k}\right) \end{cases} \tag{6.36}$$

Now we have that if $\omega^2 - k^2\omega_0^2 = \mathcal{O}(\varepsilon)$, averaging applies to (6.36).

Next we consider a specific example, the Duffing oscillator

$$f(x, \dot{x}, t) = \beta\cos(\omega t) - \delta\dot{x} - \alpha x^3, \tag{6.37}$$

with $\omega_0^2 - \omega^2 = \varepsilon\Omega$, i.e. a nearly order one resonance. The transformed system then takes the form

$$\begin{cases} \dot{u} = \dfrac{\varepsilon}{\omega}[ & \Omega\left(u\cos(\omega t) - v\sin(\omega t)\right) - \omega\delta\left(u\sin(\omega t) + v\cos(\omega t)\right) \\ & +\alpha\left(u\cos(\omega t) - v\sin(\omega t)\right)^3 - \beta\cos(\omega t)]\sin(\omega t) \\ \dot{v} = \dfrac{\varepsilon}{\omega}[ & \Omega\left(u\cos(\omega t) - v\sin(\omega t)\right) - \omega\delta\left(u\sin(\omega t) + v\cos(\omega t)\right) \\ & +\alpha\left(u\cos(\omega t) - v\sin(\omega t)\right)^3 - \beta\cos(\omega t)]\cos(\omega t). \end{cases} \tag{6.38}$$

Averaging over one period $T = \frac{2\pi}{\omega}$ we find

$$\frac{1}{T}\int_0^T \sin(\omega t)\cos(\omega t)dt = \frac{1}{2T}\int_0^T \sin(2\omega t)dt = \left[-\frac{\cos(2\omega t)}{4\omega T}\right]_0^T = 0, \tag{6.39}$$

and

$$\frac{1}{T}\int_0^T \sin^2(\omega t)dt = \frac{1}{T}\int_0^T \frac{1-\cos(2\omega t)}{2}dt = \frac{1}{T}\left[\frac{1}{2} - \frac{\sin(2\omega t)}{2\omega}\right]_0^T = \frac{1}{2}. \tag{6.40}$$

Using trigonometric identities then leads to the equation

$$\begin{cases} \dot{u} = \dfrac{\varepsilon}{2\omega}\left[-\omega\delta u - \Omega v - \dfrac{3\alpha}{4}(u^2 + v^2)v\right] \\ \dot{v} = \dfrac{\varepsilon}{2\omega}\left[\Omega u - \omega\delta v + \dfrac{3\alpha}{4}(u^2 + v^2)u - \beta\right]. \end{cases} \tag{6.41}$$

This can be further simplified by using the polar coordinates $(r, \phi)$, i.e.

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} r\cos(\phi) \\ r\sin(\phi) \end{pmatrix} \implies \begin{cases} \dot{r} = \dfrac{\varepsilon}{2\omega}\left[-\omega\delta r - \beta\sin(\phi)\right] \\ r\dot{\phi} = \dfrac{\varepsilon}{2\omega}\left[\Omega r + \dfrac{3\alpha}{4}r^3 - \beta\cos(\phi)\right]. \end{cases} \tag{6.42}$$

Next, recall that $x(t) = u(t)\cos(\omega t) - v(t)\sin(\omega t) = r(t)\cos(\omega t - \phi(t))$, i.e. we have slowly varying amplitude $r(t)$ and phase $\phi(t)$. The hyperbolic fixed points of (6.42) correspond to steady, almost sinusoidal response of the original system. We obtain the *forced response curves* by fixing $\alpha,\ \beta$, and $\delta$ and plotting the fixed point $(\overline{r}, \overline{\phi})$ of (6.42) as a function of the frequency parameter $\Omega$ (or $\frac{\omega}{\omega_0}$) as in Fig. 6.7.

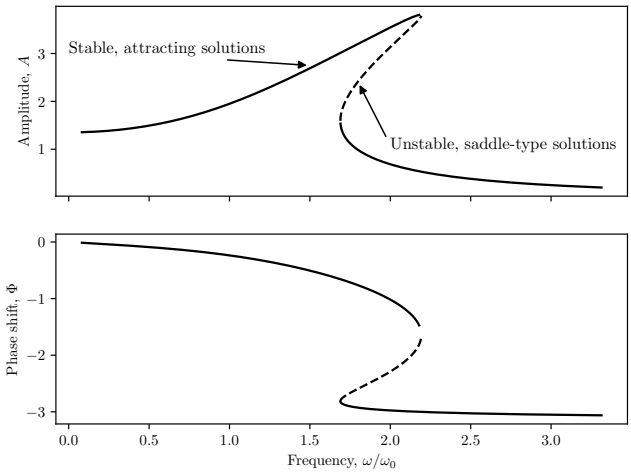


Figure 6.7: The forced response curve for the Duffing equation. The parameters are selected as $F_0 = 2, \omega_0 = 1, \alpha = 0.35, \zeta = 0.12$.

## 6.4 The Harmonic Balance Method

Averaging and other perturbation techniques are often restrictive due to their small-parameter assumptions and result in cumbersome expressions limiting their application in high-dimensional systems. To avoid these constraints, the *Harmonic Balance* method computes the Fourier series approximation of the periodic response and is formally applicable without the previous restrictions. Consider the setup

$$\dot{x} = f(x,t) = f(x,t+T); \quad x \in \mathbb{R}^n; \quad f \in \mathcal{C}_x^r,\ r \geq 1,\ f \in \mathcal{C}_t^0. \tag{6.43}$$

Now we seek a $T$-periodic response via the Fourier series.

*Remark* 6.4. It is not immediately clear that a periodic response exists in a periodically forced multi-degree of freedom mechanical system. See for instance [Breunung and Haller, 2019] for more detail.

Assume (6.43) exhibits a $T$-periodic response, say $x_p(t) = x_p(t+T)$. Then it can be expressed as a convergent Fourier series

$$x_p(t) = \sum_{k\in\mathbb{Z}} \hat{x}_k e^{ik\omega t} = x_0 + \sum_{j\in\mathbb{N}} c_j \cos(j\omega t) + s_j \sin(j\omega t). \tag{6.44}$$

Substituting this into (6.43) and using that $f$ is $T$-periodic (in both variables) and therefore admits a convergent Fourier series we find

$$\frac{d}{dt}x_p(t) = f(x_p(t), t) \tag{6.45a}$$

$$\frac{d}{dt}\sum_{k\in\mathbb{Z}} \hat{x}_k e^{ik\omega t} = \sum_{k\in\mathbb{Z}} \hat{f}_k(x_p(t), t) e^{ik\omega t} \tag{6.45b}$$

$$\sum_{k\in\mathbb{Z}} ik\omega \hat{x}_k e^{ik\omega t} = \sum_{k\in\mathbb{Z}} \hat{f}_k(x_p(t), t) e^{ik\omega t}. \tag{6.45c}$$

Now we can compare coefficients at different harmonics to find

$$ik\omega\hat{x}_k = \hat{f}_k(x_p(t), t) = \hat{f}_k\left(\sum_{j\in\mathbb{Z}} \hat{x}_j e^{ij\omega t}, t\right) \quad \forall k. \tag{6.46}$$

These are algebraic equations which give a solution for each coefficient. In general they are coupled, with a noteworthy exception if $f$ is linear. However, there are infinitely many equations to solve, so in practice we only use a finite truncation of the Fourier series to approximate the response.

$$x_H = \sum_{k=-H}^{H} \hat{x}_k e^{ik\omega t}. \tag{6.47}$$

Now we instead solve $\dot{x}_H = f_H(x_H(t), t)$, leading to only $n(2H+1)$ equations. Recall a few basic results from Fourier Analysis.

**Theorem 6.5.** *For any $F \in \mathcal{C}^0$ which is $T$-periodic, the truncation $F_H(t) = \sum_{k=-H}^{H} \hat{F}_k e^{ik\omega t}$ converges uniformly to $F(t)$ as $H \to \infty$.*

**Theorem 6.6** (Decay of Fourier coefficients)**.** *For $F \in \mathcal{C}^r$ which is $T$-periodic, the Fourier coefficients $\hat{F}_k$ decay at least as fast as $\frac{1}{|k|^{r+1}}$. For an analytic function $F \in \mathcal{C}^a$, the coefficients decay faster than any power of $k$, i. e. exponentially.*

*Remark* 6.7. Note that the decay estimate is asymptotic, hence it is generally unclear exactly where to truncate. However, the method seems to work well for many practical applications and produces fast results.

We will now apply this method in a few examples.

*Example* 6.7 (Linear spring-mass-damper oscillator). Consider the following damped linear spring

$$\ddot{x} + 2\zeta\omega_0\dot{x} + \omega_0^2 x = F(t) = F_0\cos(\omega t). \tag{6.48}$$

As only a single harmonic is sufficient to describe $F$, the harmonic coefficients are decoupled. By taking the Fourier series of $x$ and $F = \frac{1}{2}\left(F_0 e^{i\omega t} + F_0 e^{-i\omega t}\right)$ we compare coefficients

$$k = 1: \quad -\omega^2\hat{x}_1 e^{i\omega t} + i2\zeta\omega_0\omega\hat{x}_1 e^{i\omega t} + \omega_0^2\hat{x}_1 e^{i\omega t} = F_0 e^{i\omega t}; \tag{6.49a}$$

$$k = -1: \quad -\omega^2\hat{x}_{-1} e^{-i\omega t} - i2\zeta\omega_0\omega\hat{x}_{-1} e^{-i\omega t} + \omega_0^2\hat{x}_{-1} e^{i\omega t} = F_0 e^{-i\omega t}. \tag{6.49b}$$

We therefore calculate

$$\hat{x}_1 = \frac{1}{2}\left(\frac{F_0}{\omega_0^2 - \omega^2 + i2\zeta\omega_0\omega}\right); \quad \hat{x}_{-1} = \frac{1}{2}\left(\frac{F_0}{\omega_0^2 - \omega^2 - i2\zeta\omega_0\omega}\right). \tag{6.50}$$

For all other values of $k$ $\hat{x}_k$ must equal zero. Now we can calculate the famous formula from linear vibration text books for the magnitude of the response

$$|x| = \frac{F_0}{\omega_0^2\sqrt{\left(1 - \left(\frac{\omega}{\omega_0}\right)^2\right)^2 + \left(2\zeta\frac{\omega}{\omega_0}\right)^2}}. \tag{6.51}$$

*Example* 6.8 (Duffing Oscillator). Recall the Duffing oscillator

$$\ddot{x} + 2\zeta\omega_0\dot{x} + \omega_0^2 x + \alpha x^3 = F_0\cos(\omega t). \tag{6.52}$$

We consider the Ansatz

$$q_1(t) = c_1\cos(\omega t) + s_1\sin(\omega t) \implies \ddot{q}_1 + 2\zeta\omega_0\dot{q}_1 + \omega_0^2 q_1 + \alpha q_1^3 = F_0\cos(\omega t). \tag{6.53}$$

Now calculate $q_1^3$ using multi-angle trigonometric identities

$$q_1^3 = \left(c_1\cos(\omega t) + s_1\sin(\omega t)\right)^3 \tag{6.54a}$$

$$= c_1^3\cos^3(\omega t) + 3c_1^2 s_1\cos^2(\omega t)\sin(\omega t) + 3c_1 s_1^2\cos(\omega t)\sin^2(\omega t) + s_1^3\sin^3(\omega t) \tag{6.54b}$$

$$= \frac{1}{4}\left(3c_1^3 + 3c_1 s_1^2\right)\cos(\omega t) + \frac{1}{4}\left(c_1^3 - 3c_1 s_1^2\right)\cos(3\omega t) + \frac{1}{4}\left(3c_1^2 s_1 + 3s_1^3\right)\sin(\omega t) \tag{6.54c}$$

$$+ \frac{1}{4}\left(3c_1^2 s_1 - s_1^3\right)\sin(3\omega t). \tag{6.54d}$$

Now we neglect the harmonics higher than 1 and balance the remaining

$$0 = \left[\left(\omega_0^2 - \omega^2\right) c_1 + 2\zeta\omega_0\omega s_1 + \frac{3}{4}\alpha\left(c_1^3 + c_1 s_1^2\right) - F_0\right]\cos(\omega t) \tag{6.55a}$$

$$+ \left[\left(\omega_0^2 - \omega^2\right) s_1 - 2\zeta\omega_0\omega c_1 + \frac{3}{4}\alpha\left(s_1^3 + c_1^2 s_1\right)\right]\sin(\omega t) + \text{higher order harmonics.} \tag{6.55b}$$

Therefore we have two algebraic equations in $c_1$ and $s_1$. Using the polar coordinates $c_1 = A\cos(\phi)$ and $s_1 = A\sin(\phi)$ we find

$$\left(\omega_0^2 - \omega^2\right) A\cos(\phi) + 2\zeta\omega_0\omega A\sin(\phi) + \frac{3}{4}\alpha\left(A^3\cos^3(\phi) + A^3\cos(\phi)\sin^2(\phi)\right) = F_0 \tag{6.56a}$$

$$\left(\omega_0^2 - \omega^2\right) A\sin(\phi) + 2\zeta\omega_0\omega A\cos(\phi) + \frac{3}{4}\alpha\left(A^3\sin^3(\phi) + A^3\sin(\phi)\cos^2(\phi)\right) = 0. \tag{6.56b}$$

Adding $\cos(\phi)$ times the first equation to $\sin(\phi)$ times the second we obtain

$$\left(\omega_0^2 - \omega^2\right) A + \frac{3}{4}\alpha A^3 = F_0\cos(\phi). \tag{6.57}$$

Similarly, adding $\sin(\phi)$ times the first equation to $\cos(\phi)$ times the second we obtain

$$2\zeta\omega_0\omega A = F_0\sin(\phi). \tag{6.58}$$

We now take the sum of the squares of these two new equations, which yields

$$A^2\left[\left(\omega_0^2 - \omega^2 + \frac{3}{4}\alpha A^2\right)^2 + 4\zeta^2\omega_0^2\omega^2\right] = F_0^2. \tag{6.59}$$

This equation now allows us plot the amplitude $A$ versus the frequency $\omega$ as in Fig. 6.8.

Unfortunately, in contrast to averaging, harmonic balance by itself does not return any stability information. We can, however, compute the Floquet multipliers of the periodic orbit by directly solving the equation of variations.

*Example* 6.9 (Unforced and undamped Duffing oscillator). For unforced systems harmonic balance can also be used to obtain internally parameterized periodic orbits. However, the expected period is a priori unknown. Consider the following

$$\ddot{x} + \omega_0^2 x + \alpha x^3 = 0. \tag{6.60}$$

Since the system in autonomous, the phase of the response is irrelevant. Thus the simplest Ansatz for a single harmonic response is

$$x_1(t) = A\cos(\omega t), \tag{6.61}$$

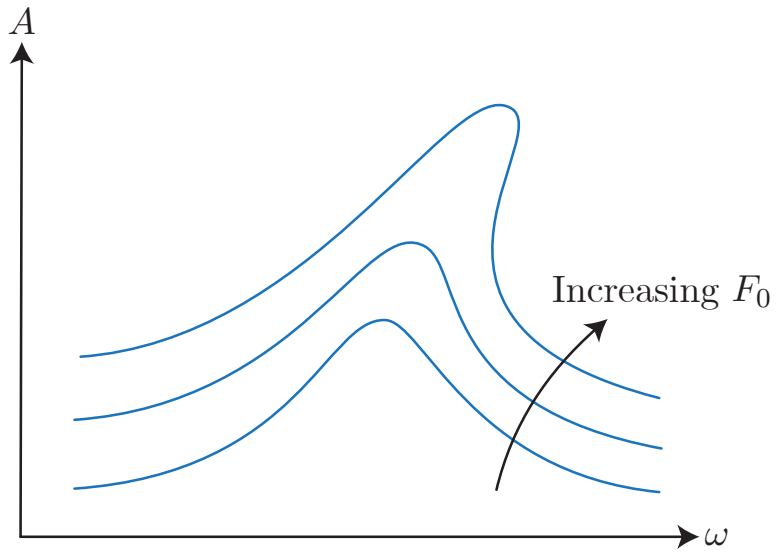


Figure 6.8: Forced response curves derived via harmonic balance for the Duffing oscillator.

for an unknown $\omega$. Applying the procedure as before and neglecting the higher harmonics we find

$$\omega^2 = \omega_0 + \frac{3}{4}\alpha A^2. \tag{6.62}$$

For unforced, conservative, systems the frequency-amplitude relation gives the so-called *backbone curve*, illustrated in Fig. 6.9.

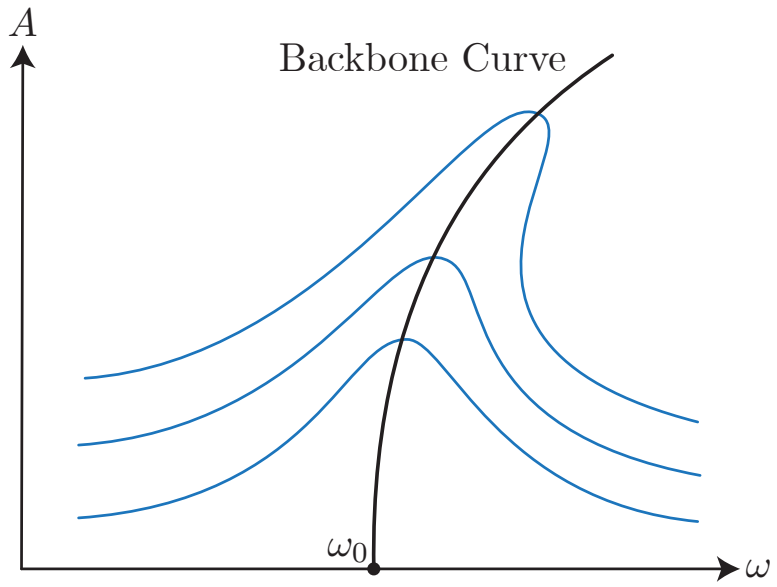


Figure 6.9: The backbone curve for the unforced and undamped Duffing oscillator. Under small forcing and damping, the backbone curve has been observed to connect the forced response curves' peaks.

# Part II

# Nonlinear Dynamics and Chaos 2

# Chapter 7

# Indroduction to chaotic dynamics

We begin with a few motivating examples.

*Example* 7.1 (Periodically forced slender beam). A beam is hanging on the inside of a rectangular frame, attached to the upper edge. Two permanent magnets are attached to the lower edge. Furthermore, there is a $T = 2\pi$-periodic forcing in the horizontal direction to the frame. The deflection of the beam is measured by the variable $x$. The setup is illustrated in Fig. 7.1. This leads to the equation of motion

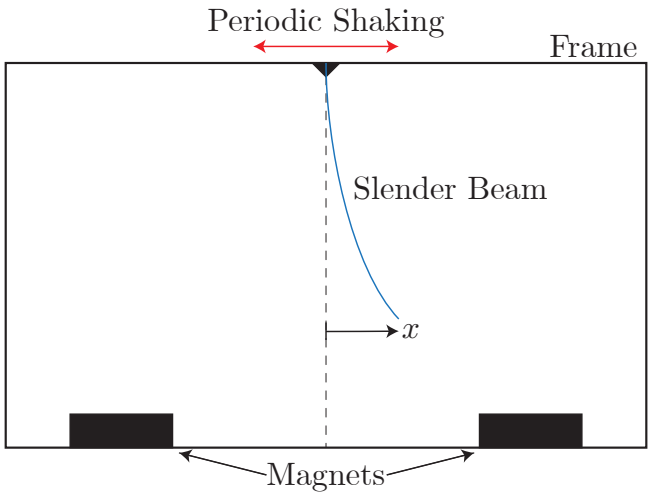


Figure 7.1: Depiction of the periodically forced slender beam. The deflection is given by $x$ and the period is $2\pi$. The two permanent magnets are represented by the black boxes.

$$\ddot{x} + \dot{x} - x + x^3 = \varepsilon \cos(t); \quad 0 \leq \varepsilon \ll 1. \tag{7.1}$$

Therefore we have a perturbed Duffing oscillator. We transform the system into a first order ODE

$$\begin{cases} \dot{x} = y \\ \dot{y} = -y + x - x^3 + \varepsilon \cos(t). \end{cases} \tag{7.2}$$

For $\varepsilon = 0$ driving and no damping, two homoclinic orbits arise for the three fixed points. However, for nonzero driving, we get seemingly chaotic behavior. Both of these regimes are depicted in Fig. 7.2.

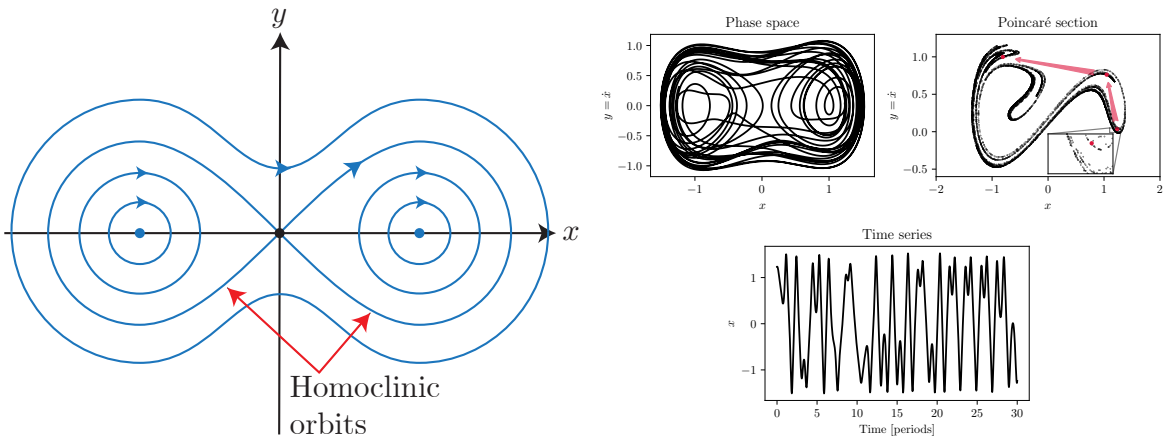


Figure 7.2: Left: Duffing oscillator with $\varepsilon = 0$ driving and no damping, the red arrows designate the homoclinic orbits. Right: (Clockwise from top left) The phase space of the driven Duffing oscillator; The Poincaré section of the driven Duffing oscillator, with the Poincaré map illustrated by the black arrows; The value of $x(t)$ over time, with no apparent pattern.

*Example* 7.2 (2-dimensional Rayleigh-Bénard convection). A fluid is held between two plates. The upper plate is cold and the lower plate is hot. This causes convections to start as hotter fluid rises and colder fluid falls. So called convection cells then form for low Rayleigh numbers. This process is illustrated in Fig. 7.3.

If the Rayleigh number $R_a$ exceeds a critical value $R_{a_{\text{crit}}}$, we have a time periodic perturbation to the velocity field. The fluid trajectories have the following equations of motion

$$\begin{cases} \dot{x} = u(x,y) + \varepsilon u_1(x,y,t) \\ \dot{y} = v(x,y) + \varepsilon v_1(x,y,t) \end{cases}; \quad \varepsilon > 0. \tag{7.3}$$

Here, the functions $u_1$ and $v_1$ are $T$-periodic. The chaotic nature of this system is shown in Fig. 7.4.

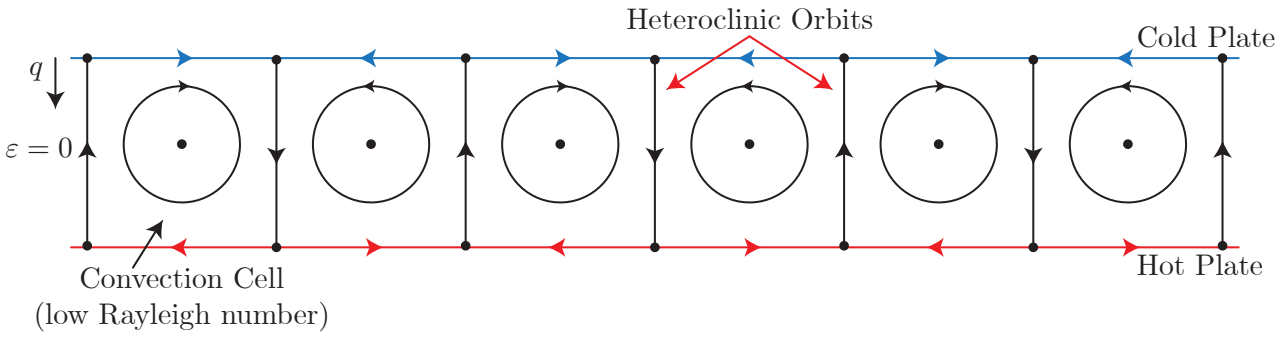


Figure 7.3: Convection cells forming between two plates of different temperature. The cold plate (blue) is vertically above the hot plate (red). Each black dot represents a fixed point of the system, with heteroclinic orbits connecting them (red arrows). This is in an unperturbed setting ($\varepsilon = 0$).

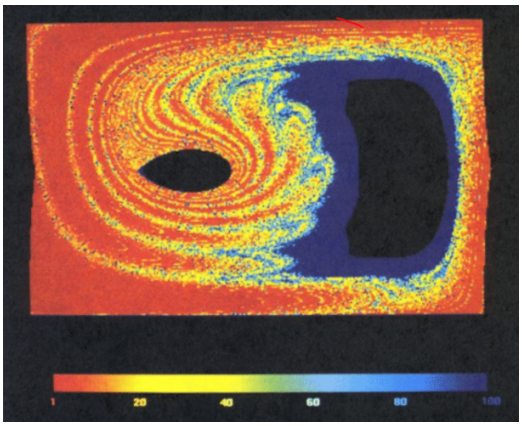

Figure 7.4: Escape time from one convection cell. The number of iterations of the Poincaré map needed for escape is given by the color scale. Here $0 < \varepsilon \leq 1$.

Moving forward we will have a common mathematical setting

$$\dot{x} = f(x) + \varepsilon g(x,t); \quad x \in \mathbb{R}^2, \quad g(x,t) = g(x,t+T); \quad 0 \leq \varepsilon \ll 1; \quad f, g \in \mathcal{C}^1. \tag{7.4}$$

We have a small $T$-periodic perturbation of a planar ODE which can be studied via a Poincaré map $P_\varepsilon^{t_0} : x_0 \mapsto x(t_0 + T; t_0, x_0)$. Assume that for $\varepsilon = 0$ the system (7.4) has a saddle type fixed point with a homoclinic (or heteroclinic) orbit $x^0(t - t_0)$. Such a system is depicted in Fig. 7.5.

*Remark* 7.1. The fixed point of $P_0^{t_0}$ given by $p_0$ is hyperbolic

$$P_0^{t_0}(p_0) = p_0; \quad DP_0^{t_0}(p_0) \text{ has eigenvalues } \lambda_1, \lambda_2 : \ |\lambda_1| < 1, \ |\lambda_2| > 1. \tag{7.5}$$

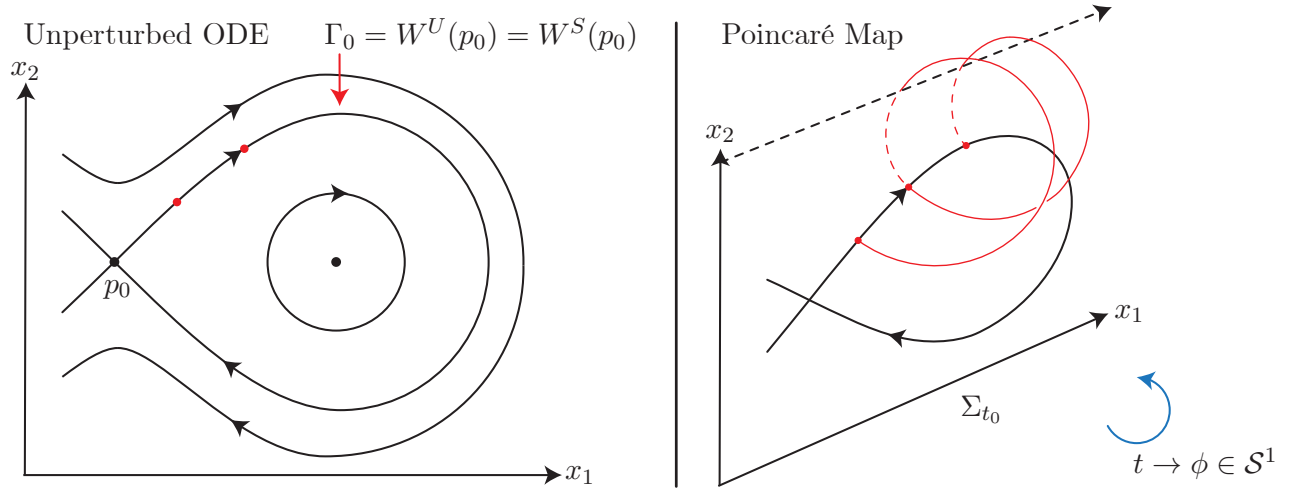


Figure 7.5: Left: An example of a system following the assumptions. The homoclinic orbit $\Gamma_0$ is equal to the local unstable and local stable manifolds. Right: The Poincaré map for $\varepsilon = 0$, with time running counterclockwise about the $x_1$ axis with period $T$.

More generally we define a hyperbolic fixed point for a map.

**Definition 7.1.** For a map $F : \mathbb{R}^n \to \mathbb{R}^n$ and a dynamical system with $x_{k+1} = F(x_k)$, the fixed point $p_0$ (i.e. $F(p_0) = p_0$) is *hyperbolic* if the linearization's $DF(p_0) \in \mathbb{R}^{n\times n}$ eigenvalues $\lambda_1, \ldots, \lambda_n$ never have unitary length: $|\lambda_i| \neq 1$ for $i = 1, \ldots, n$.

For each eigenvalue $\lambda_i$ of the linearization the corresponding eigenvector is given by $s_i$. Assume that $DF(p_0)$ is semisimple, i.e. all of the eigenvectors are linearly independent. The linearized dynamics at $p_0$ for $y = x - p_0$ small is

$$y_{k+1} = DF(p_0)y_k \implies y_k = \lambda_1^k c_1 s_1 + \ldots + \lambda_n^k c_n s_n. \tag{7.6}$$

If all of the eigenvalues have less than unitary magnitude, $|\lambda_i| < 1$ for $i = 1, \ldots, n$, then $p_0$ is asymptotically stable. Otherwise if there exists an eigenvalue with modulus strictly larger than 1, $p_0$ is unstable. The relationship of the nonlinear and linearized dynamics are shown in Fig. 7.6.

The fixed point $p_0$ along with the stable and unstable manifolds ($W^S(p_0)$ and $W^U(p_0)$) persist smoothly under small smooth perturbations to $F$ near $p_0$.

## 7.1 Consequences of hyperbolicity

Hyperbolicity has a few important consequences in our setting. Foremost, the perturbed hyperbolic fixed point $p_\varepsilon^{t_0}$ translates to a hyperbolic periodic orbit for the ODE. Furthermore, the

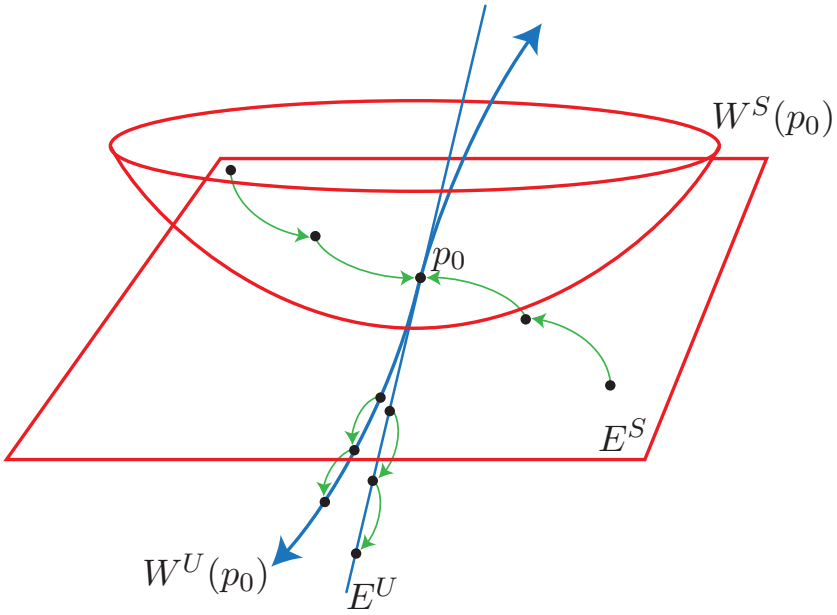


Figure 7.6: The stable ($W^S(p_0)$) and unstable ($W^U(p_0)$) manifolds drawn in relation to the stable ($E^S$) and unstable ($E^U$) subspaces. The subspaces are given as the span of the stable (resp. unstable) eigenvectors. The green arrows signify steps of the linearized (resp. original) system.

solutions of the ODE are smooth in $\varepsilon$, thus the solutions within $W^U(p_\varepsilon^{t_0})$ and $W^S(p_\varepsilon^{t_0})$ remain $\mathcal{O}(\varepsilon)$ close to $\Gamma_0$, i.e.

$$x_\varepsilon^S(t;t_0) = x^0(t-t_0) + \varepsilon a^S(t) + \mathcal{O}(\varepsilon^2); \quad t \in [t_0, \infty) \tag{7.7a}$$
$$x_\varepsilon^U(t;t_0) = x^0(t-t_0) + \varepsilon a^U(t) + \mathcal{O}(\varepsilon^2); \quad t \in (-\infty, t_0]. \tag{7.7b}$$

Now we examine what the global shapes of these manifolds are for $\varepsilon > 0$ and if they interact. To do this we will follow an idea from Poincaré, Arnold, and Melnikov. To this end we define the perpendicular vector to $f$

$$f^\perp(x^0(0)) = \begin{pmatrix} -f_2(x^0(0)) \\ f_1(x^0(0)) \end{pmatrix}. \tag{7.8}$$

We would like to use this to measure the distance between $x_\varepsilon^U(t_0;t_0)$ and $x_\varepsilon^S(t_0;t_0)$. The outlook for this is shown in Fig. 7.7. Thus we have a signed distance function

$$d_\varepsilon(t_0) = \frac{\langle f^\perp(x^0(0)), x_\varepsilon^U(t_0;t_0) - x_\varepsilon^S(t_0;t_0)\rangle}{|f^\perp(x^0(0))|} = \varepsilon \frac{\langle f^\perp(x^0(0)), a^U(t_0) - a^S(t_0)\rangle}{|f^\perp(x^0(0))|} + \mathcal{O}(\varepsilon^2). \tag{7.9}$$

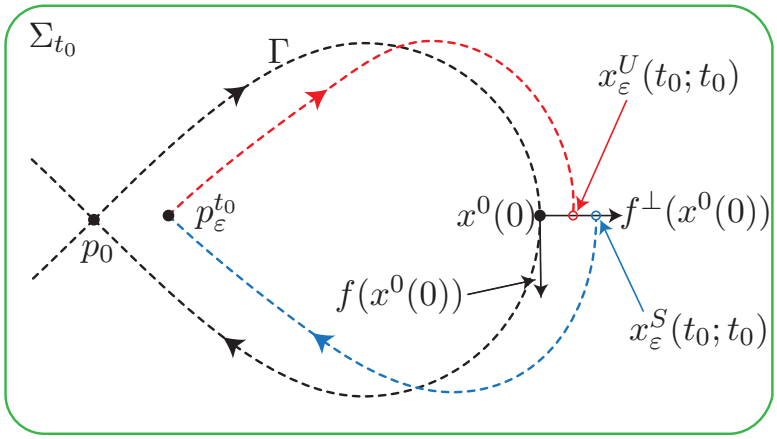


Figure 7.7: The Poincaré-Arnold-Melnikov idea. The dotted black line signifies $\Gamma$, the red dot $x_\varepsilon^U(t_0;t_0)$, the blue dot $x_\varepsilon^S(t_0;t_0)$, and the arrow pointing to the right $f^\perp$. The plane is $\Sigma_{t_0}$.

The numerator in the second equation is called the *Melnikov function*

$$M(t_0) = \left\langle f^\perp(x^0(0)), a^U(t_0) - a^S(t_0) \right\rangle . \tag{7.10}$$

In fact Melnikov proved an identity for this function.

**Theorem 7.2** (Melnikov)**.**

$$M(t_0) = \int_{-\infty}^{\infty} \langle f^\perp(x^0(t-t_0)), g(x^0(t-t_0),t)\rangle dt. \tag{7.11}$$

*For the proof, see Theorem 4.5.3 in [Guckenheimer and Holmes, 1990].*

*Remark* 7.3. The integral in Melnikov's theorem converges as $|g(x^0(t-t_0),t)|$ is globally bounded as $t \to \pm\infty$ and $|f^\perp| = |f|$ and

$$\lim_{t\to\pm\infty} \left| f(x^0(t-t_0)) \right| = 0, \tag{7.12}$$

exponentially as $p_0$ is a hyperbolic fixed point.

*Remark* 7.4. In order to evaluate $M(t_0)$ we do not need to solve the perturbed ODE $\dot{x} = f(x) + \varepsilon g(x,t)$, instead we only need $x^0(t-t_0)$.

Now observe that

$$d_\varepsilon(t_0) = 0 \iff \varepsilon \frac{M(t_0)}{|f^\perp(x^0(0))|} + \mathcal{O}(\varepsilon^2) = 0 \overset{\varepsilon\neq 0}{\iff} \underbrace{\frac{M(t_0)}{|f^\perp(x^0(0))|} + \mathcal{O}(\varepsilon) = 0}_{F(t_0,\varepsilon)=0}. \tag{7.13}$$

Now we would like to know when we can find a solution $t_0(\varepsilon)$ such that $d_\varepsilon(t_0(\varepsilon)) = 0$ for $\varepsilon > 0$. To do this we use the Implicit Function Theorem. First assume that $F(\bar{t}_0, 0) = 0$, i.e. $M(\bar{t}_0) = 0$, and $\frac{\partial F}{\partial t_0}(\bar{t}_0), 0) \neq 0$, i.e. $\frac{\partial}{\partial t_0} M(\bar{t}_0) \neq 0$. If this condition is fulfilled the root is called *transverse*. Then there exists a unique $t_0(\varepsilon) = \bar{t}_0 + \mathcal{O}(\varepsilon)$ which solves $F(t_0(\varepsilon), \varepsilon) = 0$ for $\varepsilon \neq 0$ small enough. Also $t_0(\varepsilon)$ is smooth if $F(t_0, \cdot)$ is smooth.

A transverse zero for $M(t_0)$ implies that the Melnikov distance $d_\varepsilon(t_0(\varepsilon) = 0$, in turn implying that the stable and unstable manifolds intersect, $W^S(p_\varepsilon^{t_0}) \cap W^U(p_\varepsilon^{t_0}) \neq \emptyset$. Therefore we have an element $q \in W^S(p_\varepsilon^{t_0}) \cap W^U(p_\varepsilon^{t_0})$, for this $q$ we also have that for every $n \in \mathbb{Z}$

$$P_\varepsilon^n(q) \in W^S(p_\varepsilon^{t_0}) \cap W^U(p_\varepsilon^{t_0}). \tag{7.14}$$

Therefore we have that for each iterate of the Poincaré map there is a unique point in both the stable and unstable manifolds, so these must intersect each other infinitely many times. This behavior is shown in Fig. 7.8. In the homoclinic tangle, we can see the accumulation

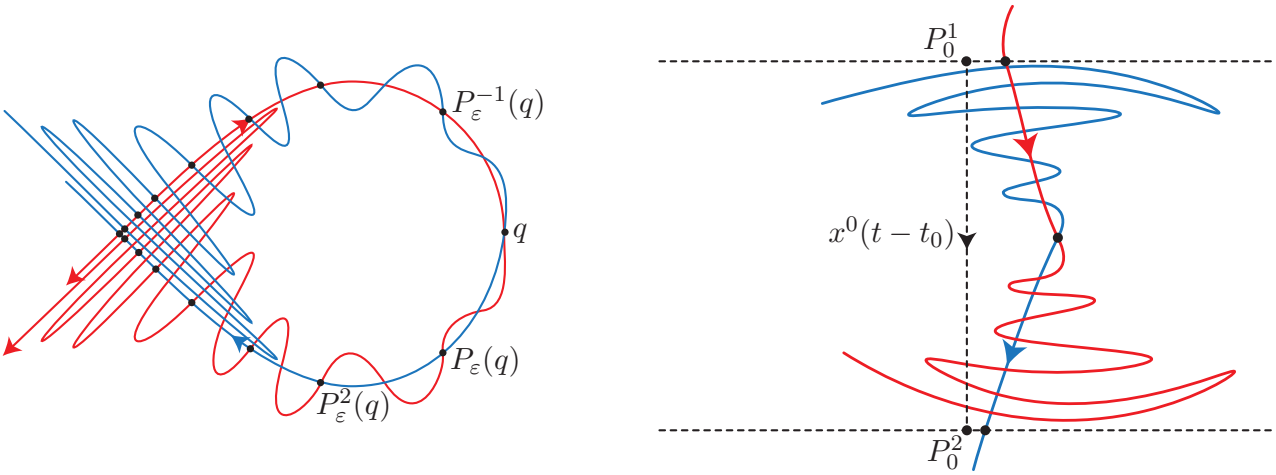


Figure 7.8: The infinite intersections of the stable and unstable manifolds for a transverse zero. Left: The homoclinic case, this is aptly called a *homoclinic tangle*. Right: The heteroclinic case.

of trajectories that are caused by the accumulation of the iterates of the Poincaré map. The accumulation can be formally seen as a consequence of the $\Lambda$-lemma [Palis, 1967, Palis, 1969].

*Remark* 7.5. It is impossible for stable and unstable manifolds to intersect themselves. This is usually argued by stating that $q$ cannot have two distinctive preimages under a diffeomorphism. However, this is not sufficient as a self intersection does not imply that two distinct preimages exist, for instance a looping manifold structure and suitable Poincaré map as in Fig. 7.9. A priori, this is possible to occur. In this case the existence of one loop actually implies the

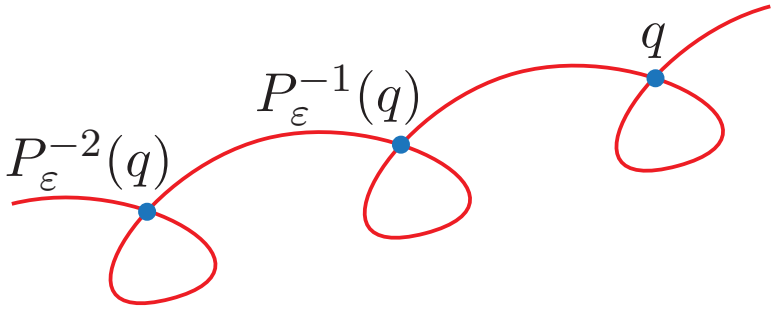


Figure 7.9: An example for a self-intersecting manifold without $q$ having multiple preimages.

existence of infinitely many converging to $p_\varepsilon$. Thereby there must exist a loop in every arbitrarily small neighborhood of $p_\varepsilon$, this then contradicts the Hartman-Grobman Theorem as our system must be topologically equivalent to the linear saddle in a small enough neighborhood.

*Remark* 7.6. For every pair of intersections $P_\varepsilon^k(q)$ and $P_\varepsilon^{k+1}(q)$, there exists at least another intersection between the stable and unstable manifolds. This is due to the fact that $P_\varepsilon$ is orientation preserving (i.e. $DP_\varepsilon(q)$ is an orientation preserving linear map). The implication here is illustrated in Fig. 7.10.

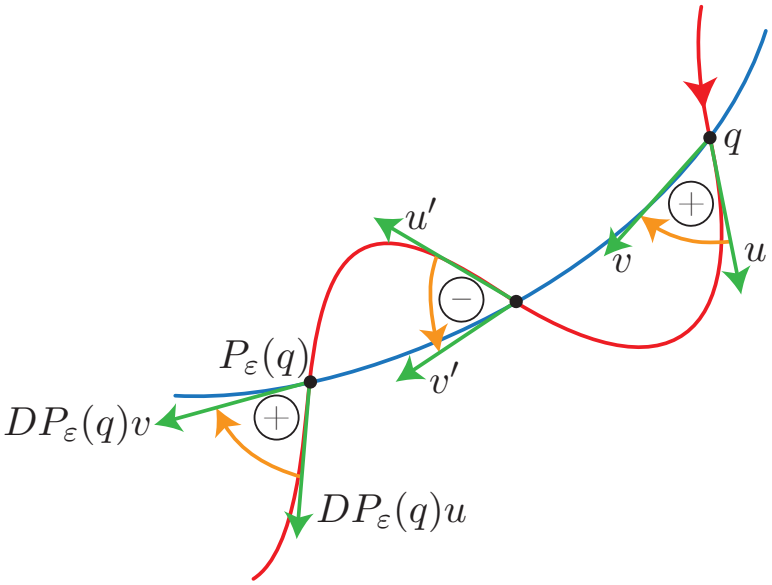


Figure 7.10: Another intersection of the stable and unstable manifolds exists between two sequential iterates of the Poincaré map due to the preservation of orientation under this map.

The orientation preserving nature of $P_\varepsilon$ can be seen by using Liouville's Theorem. It states that for a general dynamical system $\dot{x} = f(x,t)$, the linearization of its flow-map $F_{t_0}^t : x_0 \mapsto$

$x(t; t_0, x_0)$ satisfies

$$\det(DF_{t_0}^t(p)) = \exp\left(\int_{t_0}^t \nabla \cdot (f(x(s; t_0, p), s))ds\right). \tag{7.15}$$

In our case we find

$$\det(DP_\varepsilon(q)) = \exp\left(\int_{t_0}^{t_0+T} \nabla \cdot (f + \varepsilon g)|_{x_\varepsilon(t);\ x_\varepsilon(t_0)=q})ds\right) > 0 \tag{7.16}$$

Therefore $DP_\varepsilon(q)$ is orientation preserving, in fact Poincaré maps in general are orientation preserving.

*Example* 7.3 (The forced-damped Duffing equation). Recall the dynamical system

$$\begin{cases} \dot{x} = y \\ \dot{y} = x - x^3 + \varepsilon(\gamma\cos(\omega t) - \delta y) \end{cases}; \quad |\varepsilon| \ll 1. \tag{7.17}$$

Here the forcing amplitude is given by $\gamma$ and the linear damping coefficient by $\delta$. We separate the right hand side into two parts

$$f(x, y) = \begin{pmatrix} y \\ x - x^3 \end{pmatrix}, \quad g(x, y) = \begin{pmatrix} 0 \\ \gamma\cos(\omega t) - \delta y \end{pmatrix}. \tag{7.18}$$

The Hamiltonian for the unperturbed system is given by

$$H = \frac{1}{2}y^2 - \frac{1}{2}x^2 + \frac{1}{4}x^4. \tag{7.19}$$

Since the Hamiltonian is conserved, the trajectories of the unperturbed system are contained in the level sets of $H$. This allows us to draw the phase portrait that is shown in Fig. 7.11. The homoclinic orbit $\Gamma_0$ is contained in the $H = 0$ level set of the Hamiltonian, which means that along this orbit we have

$$\frac{1}{2}y^2 - \frac{1}{2}x^2 + \frac{1}{4}x^4 = 0. \tag{7.20a}$$

$$\frac{dx}{dt} = y = \pm\sqrt{x^2 - \frac{1}{2}x^2} = F(x). \tag{7.20b}$$

Locally we can obtain a parametrization of the heteroclinic orbit by solving this separable differential equation as

$$\frac{dx}{dt} = F(x) \implies \int_0^x \frac{d\tilde{x}}{F(\tilde{x})} = \int_0^t d\tilde{t}. \tag{7.21}$$

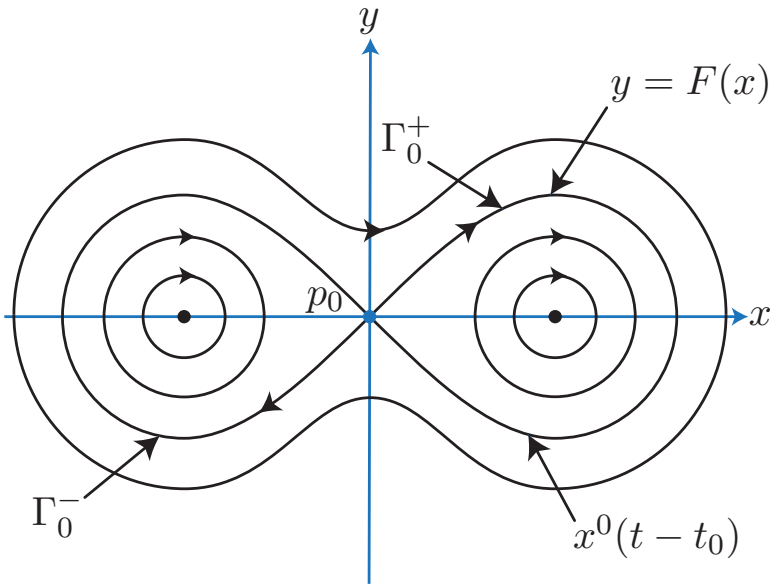


Figure 7.11: The phase portrait for the forced-damped Duffing oscillator for $\varepsilon = 0$.

Hence, on $\Gamma_0^+$, we have

$$x^0(t) = \begin{pmatrix} \sqrt{2}\text{sech}(t) \\ -\sqrt{2}\text{sech}(t)\tanh(t) \end{pmatrix}. \tag{7.22}$$

As discussed above, for $\varepsilon > 0$, the homoclinic orbit splits into a stable and an unstable manifold. We can now calculate the Melnikov function to study whether these manifolds intersect.

$$M^+(t_0) = \int_{-\infty}^{\infty} \langle f^\perp(x^0(t-t_0), g(x^0(t-t_0), t)\rangle dt \tag{7.23a}$$

$$= -\frac{4\delta}{3} + \sqrt{2}\gamma\pi\omega\text{sech}\left(\frac{\pi\omega}{2}\right)\sin(\omega t_0). \tag{7.23b}$$

Next, we ask if the Melnikov function is equal to zero. For an accumulation of intersection points, a homoclinic tangle, we also need a transverse zero. To this end we define $R^0(\omega) = \frac{4\cosh\left(\frac{\pi\omega}{2}\right)}{3\sqrt{2}\pi\omega}$. Thus the zeros of the Melnikov function are given by

$$R^0(\omega) = \frac{\gamma}{\delta}\sin(\omega t_0). \tag{7.24}$$

For a zero to exist, $R^0(\omega)$ cannot exceed the value $\frac{\gamma}{\delta}$, as $\sin(\omega t_0)$ is bounded by 1. Therefore when we have $R^0(\omega) \leq \frac{\gamma}{\delta}$, a root of the Melnikov function exists. This root is non-transverse only when it is quadratic, i.e. when only a peak of the sine function intersects with $R^0(\omega)$. This

occurs exactly when $R^0(\omega) = \frac{\gamma}{\delta}$. In all other cases, the sine wave intersects the curve of $R^0(\omega)$ transversely, i.e. the zero is transverse. This is illustrated in Fig. 7.12, where equality occurs along the graph of $R^0$, that is the set on which non-transverse zeros exist. The area above this curve contains all of the point for which the zero is transverse.

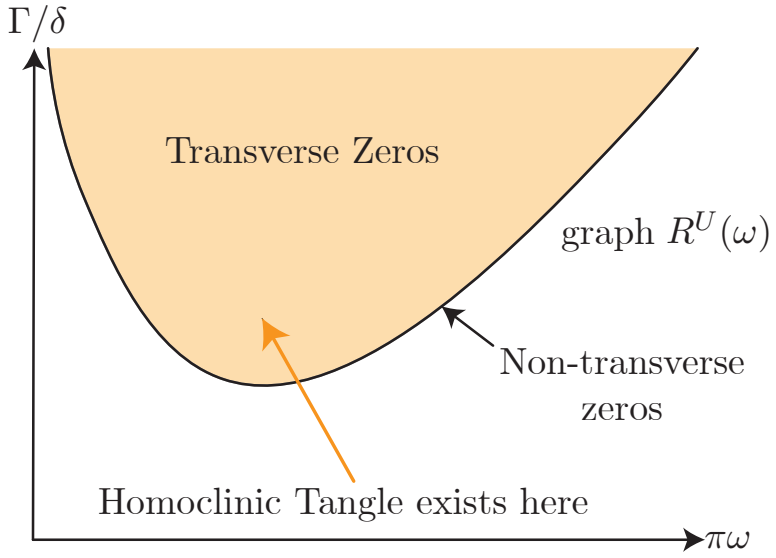


Figure 7.12: The yellow line designating the area above the curve where homoclinic tangles exist for the forced-damped Duffing oscillator.

## 7.2 Dynamics near the homoclinic tangle & Smale's horseshoe map

We now continue and examine the dynamics near the homoclinic tangle. Consider the dynamical system

$$\dot{x} = f(x) + \varepsilon g(x,t); \quad x \in \mathbb{R}^2; \quad f, g \in \mathcal{C}^1; \quad g(x,t) = g(x,t+T). \tag{7.25}$$

With a linear change of coordinates we can align our coordinate axes with the stable and unstable subspaces of the fixed point. Near the fixed point this leads to *Smale's construction*, which is as shown as in Fig. 7.13.

A model for the right panel of Fig. 7.13 is given by *Smale's horseshoe map*. This is defined by first considering the unit square $S = [0,1]^2 \subset \mathbb{R}^2$, next the map itself is given by $f : S \to \mathbb{R}^2$. The map $f$ will not be defined explicitly, and instead geometrically. The unit square is divided

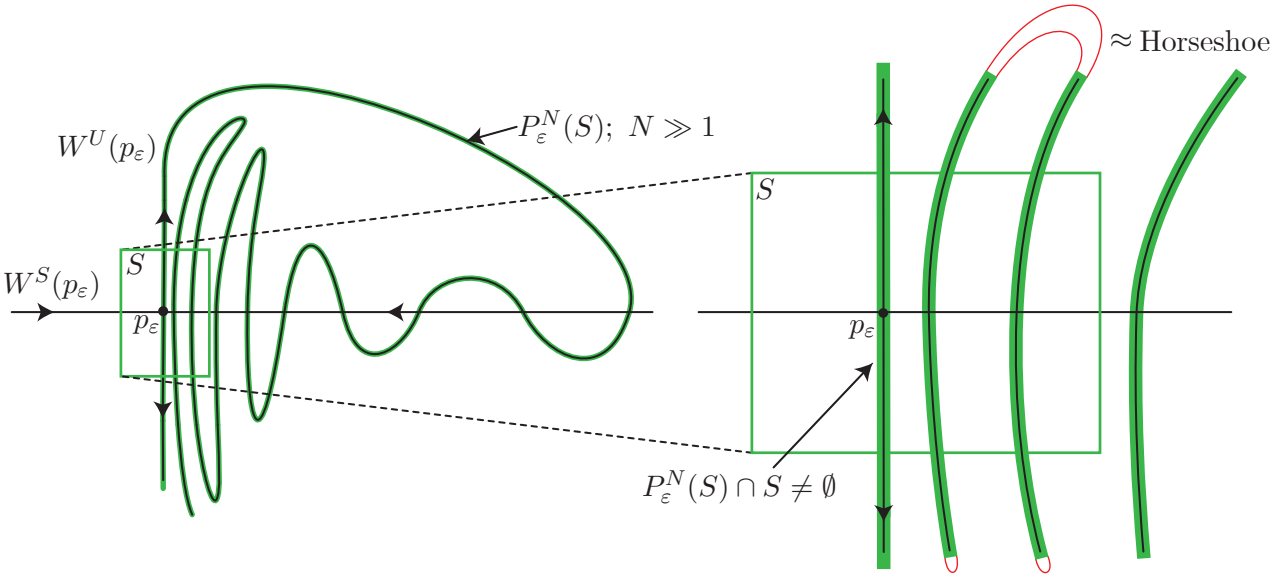


Figure 7.13: Smale's construction is found through an appropriate change of coordinates. Left: The full construction, the $N$ is large enough such that $P_\varepsilon^N(S) \cap S = \emptyset$, where $S$ is the green rectangle around $p_\varepsilon$. Right: Focus on the set $S$, the red lines designate a horseshoe-like structure.

into three horizontal strips, the upper one being labelled as $H_2$ and the lower $H_1$. The square is then squeezed horizontally and stretched vertically, and then smoothly bent to a horseshoe such that the previously upper horizontal strip now forms a vertical strip on the right side of the box $V_2$. Similarly, the lower horizontal strip now forms a vertical strip on the left $V_1$. Now the unit square can also be divided into three vertical strips, on the left $V_1$, an area in the middle, and on the right $V_2$. What was previously part of the middle horizontal strip is now mapped outside the unit square. Because the map is continuous, these points connect $V_1$ and $V_2$, but note that precisely defining what happens to these points is not essential in Smale's construction. The horseshoe map is depicted in Fig. 7.14. This is a good model for the dynamics near the homoclinic tangle, if we consider the horseshoe map $f$ to be many iterations of the Poincaré map $P_\varepsilon^N$ (for $N \gg 1$).

Using this construction, we would like to know if any initial conditions stay in $S$ for long times (after many iterations of the horseshoe map), despite the overall instability coming from $P_\varepsilon$. To further study this we introduce a couple of definitions.

**Definition 7.2.** The set of all points that stay in $S$ under $k$ backwards iterations is given by $V^k$.

**Definition 7.3.** For a map $f : X \to Y$, we denote the *preimage* of a set $U \subset Y$ (or point

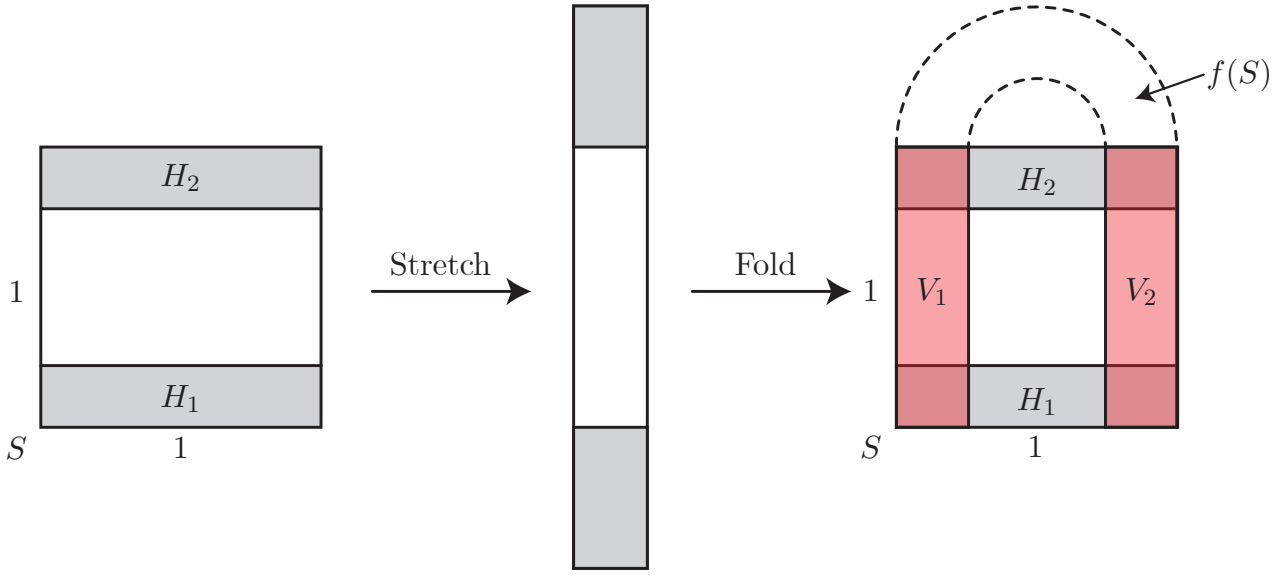


Figure 7.14: The two geometric steps of the Smale map, first a vertical stretch and then a fold into the horseshoe.

$p \in Y$) as

$$\boxed{f^{-1}(U) = \{x \in X : \ f(x) \in U\}\,.} \tag{7.26}$$

The set of points that will be mapped into $S$ in one iteration is given by $S \cap f^{-1}(S)$. This means, that the set of points that stay in $S$ for one backward iteration, $V^1$, is

$$V^1 = f(S \cap f^{-1}(S)) = f(S) \cap f(f^{-1}(S)) = f(S) \cap S. \tag{7.27}$$

However, from the definition of the horseshoe map, $V^1 = f(S) \cap S$ is $V_1 \cup V_2$, which is a closed, nonempty set. Similarly, for 2 backward iterations we have

$$V^2 = f^2\left(S \cap f^{-1}(S) \cap f^{-2}(S)\right) = f^2(S) \cap f(S) \cap S = V_{11} \cup V_{12} \cup V_{22} \cup V_{21} = \bigcup_{i_k \in \{1,2\}} V_{i_1 i_2}. \tag{7.28}$$

The sets $V_{ij}$ are defined in Fig. 7.15, the indices for each component are given by their orbit: The first index denotes which vertical strip $V_i$ ($i \in \{1, 2\}$) the preimage of a point is in. The second index denotes which vertical strip is it mapped into. That is, $V_{11}$ is the set of points that are mapped from $V_1$ to $V_1$ under one iteration, $V_{12}$ is the set of points mapped from $V_1$ to $V_2$ and so on. Also the set $V^2$ is closed, nonempty, contained in $V^1$, and has $2^2$ components.

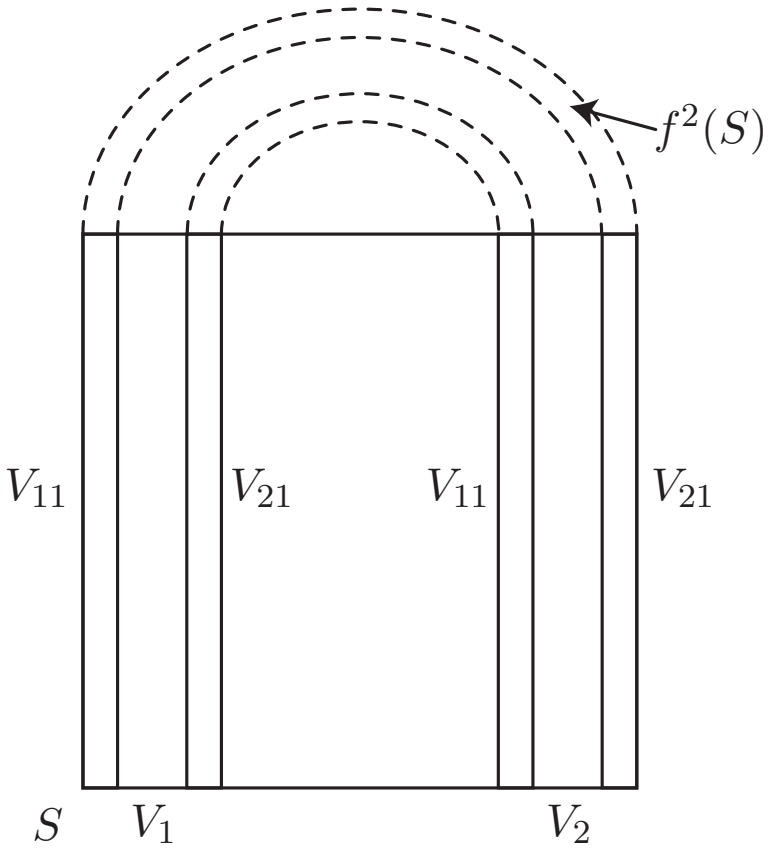


Figure 7.15: The four components of $V^2$ with their index labels.

This continues for any number of backward iterates $k$

$$V^k = f^k\left(S \cap f^{-1}(S) \cap \ldots \cap f^k(S)\right) = f^k(S) \cap \ldots \cap f^1(S) \cap S = \bigcup_{i_j \in \{1,2\}} V_{i_1 \ldots i_k}. \tag{7.29}$$

This set has $2^k$ disjoint components (vertical strips) and is closed, nonempty, and contained in $V^{k-1}$. Therefore we have an infinite, nested sequence of nonempty and closed sets, therefore by Cantor's theorem we have

$$\boxed{V^\infty = \bigcap_{i=1}^{\infty} V^i \neq \emptyset.} \tag{7.30}$$

*Remark* 7.7. Cantor's theorem comes from the famous Cantor dust which arises by first removing the open middle third of the unit interval in $\mathbb{R}$, and then iteratively removing the open middle third of each remaining interval. The Cantor dust is a nonempty (in fact uncountably

infinite) and disconnected set of points. For more on Cantor's theorem and Cantor's dust, see any book on measure theory.

**Definition 7.4.** A set $U$ is called *perfect* if it is closed and has no isolated points, i.e. given any point $x \in U$ and any distance $\varepsilon$ we can find a distinct element $y \in U$ such that $\|x-y\| < \varepsilon$.

*Remark* 7.8. This definition uses that the topology on $U$ was induced by a metric, which may not be the case in general, but is not relevant to the content here.

In our case $V^\infty$ is in fact a *Cantor set of lines*, i.e. it is a closed, nonempty set which is totally disconnected (the largest connected component is a line), and perfect.

*Remark* 7.9. By a Cantor set of lines, we mean that there is a continuous, surjective map from this set of lines to a Cantor set of points. This map could be the one taking points along a vertical line to an intersection of the line with an arbitrarily selected horizontal line.

**Definition 7.5.** Analogous to before we define the set of all points that stay in $S$ under $k$ forwards iterations as $H^k$. For each $k$ the set $H^k$ is a nonempty set of $2^k$ closed, disjoint, horizontal strips, forming a nested sequence.

By similar reasoning as before we have

$$\boxed{H^\infty = \bigcap_{k=1}^{\infty} H^k \neq \emptyset.} \tag{7.31}$$

**Definition 7.6.** The set of all points that stay in $S$ for all forward and backward iterates of the Smale horseshoe map is given by

$$\boxed{\Lambda = H^\infty \cap V^\infty.} \tag{7.32}$$

The geometry of this set is shown in Fig. 7.16.

*Remark* 7.10. The set $\Lambda$ is a Cantor set of points.

We have then identified the invariant set $\Lambda$ for the horseshoe map. Returning to our case study of a Poincaré map with a homoclinic tangle, the points in $\Lambda$ keep coming back forever to the unstable periodic orbit as shown in Fig. 7.17.

## 7.3 Dynamics of Smale's horseshoe map on the invariant set $\Lambda$

We label any point $p \in \Lambda$ with an infinitely long binary sequence, which reflects the itinerary of $p$ in $\Lambda$ under iterations of $f$. For a given $p$, if we have

$$\ldots,\ f^{-2}(p) \in V_{s_{-3}},\ f^{-1}(p) \in V_{s_{-2}},\ p \in V_{s_{-1}},\ p \in H_{s_0},\ f(p) \in H_{s_1},\ f^2(p) \in H_{s_2},\ \ldots \tag{7.33}$$

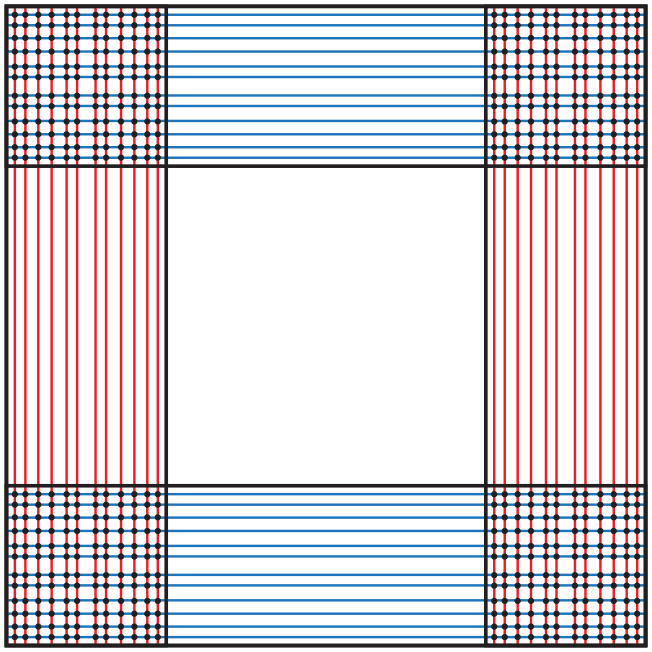

Figure 7.16: The geometry of the set $\Lambda$, with the four regions containing points circled in orange. The red vertical lines represent the set $V^\infty$ and the blue horizontal lines the set $H^\infty$.

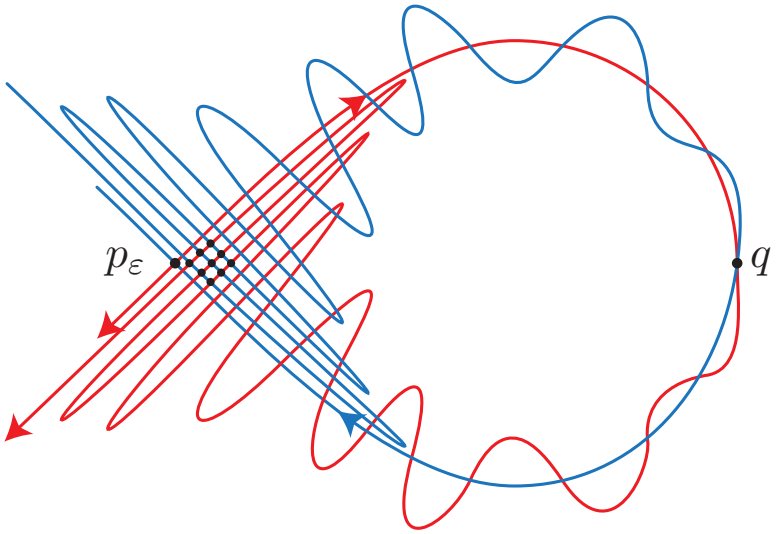


Figure 7.17: The homoclinic tangle with the elements of $\Lambda$ being denoted by the black dots within $S$, the neighborhood around $p_\varepsilon$.

for $s_i \in \{1, 2\}$ then we would have the itinerary

$$\boxed{s = \ldots s_{-3}s_{-2}s_{-1}\bullet s_0 s_1 s_2 \ldots .} \tag{7.34}$$

This is a unique symbolic encoding for all points in $\Lambda$. Therefore there is a one-to-one correspondence between points $p$ in the invariant set $\Lambda$ and sequences $s$ in the space $\Sigma$ of doubly infinite sequences of two symbols $\Sigma$. Such a space is just the space of outcomes for an infinite coin-tossing experiment. This equivalence of $\Lambda$ to the doubly infinite sequence of two symbols leads us to *symbolic dynamics*: the analogue of the horseshoe map on the symbol space $\Sigma$.

In order to study these symbolic dynamics we need to know what the symbolic encoding of $f(p)$ is. For the analogue $\ldots s_{-2}s_{-1}\bullet s_0 s_1 s_2 \ldots = s \in \Sigma$ of $p \in \Lambda$, we can read off that $f(p) \in H_{s_1}$. Continuing in this fashion, we can see that $f(f(p)) = f^2(p) \in H_{s_2}$ and so on, therefore $f(p) \in H^\infty$. By recalling that $f(H_{s_0}) = V_{s_0}$ we can see that $f(p) \in V_{s_0}$. Applying the definition of $f$ we find that $f^{-1}(f(p)) = p \in V_{s_{-1}}$. Therefore $f(p) \in V^\infty$ and thereby $f(p) \in \Lambda$.

From our examination of the forward iterates of $f(p)$ we know that the symbolic encoding of $f(p)$ to the right of the dot is

$$\bullet s_1 s_2 s_3 \ldots . \tag{7.35}$$

Similarly, from our examination of the backwards iterates of $f(p)$ we know that the symbolic encoded of $f(p)$ to the left of the dot is

$$\ldots s_{-2}s_{-1}s_0\bullet. \tag{7.36}$$

Putting these together we get the symbolic encoding for $f(p)$ as

$$\ldots s_{-2}s_{-1}s_0\bullet s_1 s_2 s_3 \ldots \in \Sigma. \tag{7.37}$$

So the encoding of $f(p)$ in the symbolic space is just a shift to the left on the encoding of $p$.

**Definition 7.7.** We define the following three maps:

(i) The *symbolic encoding map*

$$\boxed{\phi : \Lambda \to \Sigma; \quad p \mapsto s = \phi(p);} \tag{7.38}$$

This map can be shown to be a homeomorphism, i.e. it is continuous and has a continuous inverse. For a proof we refer to Theorem 5.1.1 in [Guckenheimer and Holmes, 1990].

(ii) The *Smale horseshoe map restricted to* $\Lambda$

$$\boxed{\tilde{f} : \Lambda \to \Lambda; \quad p \mapsto f(p);} \tag{7.39}$$

(iii) The *Bernoulli shift map*

$$\boxed{\sigma : \Sigma \to \Sigma; \quad \ldots s_{-2}s_{-1}\bullet s_0 s_1 s_2 \ldots \mapsto \ldots s_{-2}s_{-1}s_0\bullet s_1 s_2 s_3 \ldots .} \tag{7.40}$$

We have that the following diagram commutes

$$\begin{array}{ccc} \Sigma & \xrightarrow{\sigma} & \Sigma \\ \phi\uparrow & & \phi\uparrow \\ \Lambda & \xrightarrow{\tilde{f}} & \Lambda \end{array} . \tag{7.41}$$

Therefore $\tilde{f} = f|_\Lambda$ is topologically conjugate to a Bernoulli shift on $\Sigma$

$$\tilde{f} = \phi^{-1} \circ \sigma \circ \phi. \tag{7.42}$$

Hence, orbits of $f$ are taken to those of $\sigma$, with their orientation preserved.

*Remark* 7.11. To denote an infinitely repeating string, the bar notation will be used i.e.

$$\overline{1} = 111\ldots, \quad \overline{123} = 123123\ldots. \tag{7.43}$$

It is possible to classify all of the orbits of $\sigma$ in $\Sigma$, the classes and their cardinality are as follows:

(i) **Fixed points** There are exactly two fixed points of $\sigma$, namely $\overline{1}.\overline{1}$ and $\overline{2}.\overline{2}$;

(ii) **Period-two orbits** A single such orbit exists, that of $\overline{12}.\overline{12}$, as $\sigma^2(\overline{12}.\overline{12}) = \sigma(\overline{21}.\overline{21} = \overline{12}.\overline{12}$;

(iii) **Period-three orbits** There exists two such orbits, namely $\overline{112}.\overline{112} \to \overline{121}.\overline{121} \to \overline{211}.\overline{211}$ and $\overline{221}.\overline{221} \to \overline{212}.\overline{212} \to \overline{122}.\overline{122}$.

(iv) **Period-$k$ orbits** For every natural number $k$, there exists

$$\boxed{N(k) = \frac{1}{k}\left(2^k - \sum_{\langle i,k\rangle} iN(i)\right)} \tag{7.44}$$

orbits of period exactly $k$, the notation $\langle i, k\rangle$ signifies integers $i$ which are divisors of $k$.

There exists a countable infinity (i.e. the elements can be organized uniquely into an infinite countable sequence) of periodic orbits of arbitrarily high period. The countable sequence ordering can be done by listing each orbit in the row corresponding to its minimal period. An orbit of period $k$ is made up of $k$ points along that orbit. We assign a single number to the orbit by concatenating the decimal representations of the individual points along the orbit, starting from the one with the smallest value. Within a row, we can then order the periodic orbits based on the magnitudes of the concatenated numbers representing them. Thus $\overline{112}.\overline{112}$

would go in the third row as it has the minimal period 3, as the fixed points have period length 0 and occupy the first row, and in the leftmost position as $\overline{112} \prec \overline{122}$. Each row has a finite number of elements, so ordering each row is not a problem. Then the complete order is done by starting with 0 for the first element in the first row (the fixed point $\overline{1}.\overline{1}$), then reading from left to right before continuing to the next row. This list is well defined and countably infinite, containing all of the orbits of $\sigma$.

Before further study, it is important to introduce a notion of distance on the space $\Sigma$.

**Definition 7.8.** The distance on the space $\Sigma$ between two elements $s$ and $s'$ is given by

$$d(s, s') = \sum_{i=-\infty}^{\infty} \frac{|s_i - s'_i|}{2^{|i|}}. \tag{7.45}$$

Therefore two symbols are close if they agree on large central blocks.

*Remark* 7.12. The function $d$ satisfies all the axioms of a distance and the tuple $(\Sigma, d)$ forms a complete metric space.

**Proposition 7.13** (The periodic orbits are dense)**.** *The set of all periodic orbits is dense in* $\Sigma$.

**Proposition 7.14.** *Consider a* $k$*-periodic point* $s = \overline{a}.\overline{a}$*, where* $a$ *is a* $k$*-long binary sequence. Let* $s' = \overline{a}a'.\overline{a}$*, with* $a' \neq a$*, but an otherwise arbitrary finite length string. Clearly* $s \neq s'$*, therefore* $d(s, s') \neq 0$*, when we take the limit we find*

$$\lim_{n \to \pm\infty} d(\sigma^n(s), \sigma^n(s')) = 0. \tag{7.46}$$

Thus the underlying periodic orbit is unstable. Furthermore, $s$ has a countable infinity of homoclinic orbits. As the periodic orbit was arbitrary, these statements hold for all periodic orbits, hence all periodic orbits are unstable and have a countable infinity of homoclinic orbits.

*Remark* 7.15. Also note that any two periodic orbits (of arbitrary period) are connected by a countable infinity of heteroclinic orbits.

Together, there exists a countable infinity of periodic and asymptotically periodic orbits, a sketch of this behavior is shown in Fig. 7.18.

**Definition 7.9.** A point $s$ is asymptotically periodic if there is a periodic orbit $p$ such that

$$\boxed{\lim_{n \to \infty} d(\sigma^n(s), \sigma^n(p)) = 0.} \tag{7.47}$$

For example, heteroclinic orbits are asymptotically periodic.

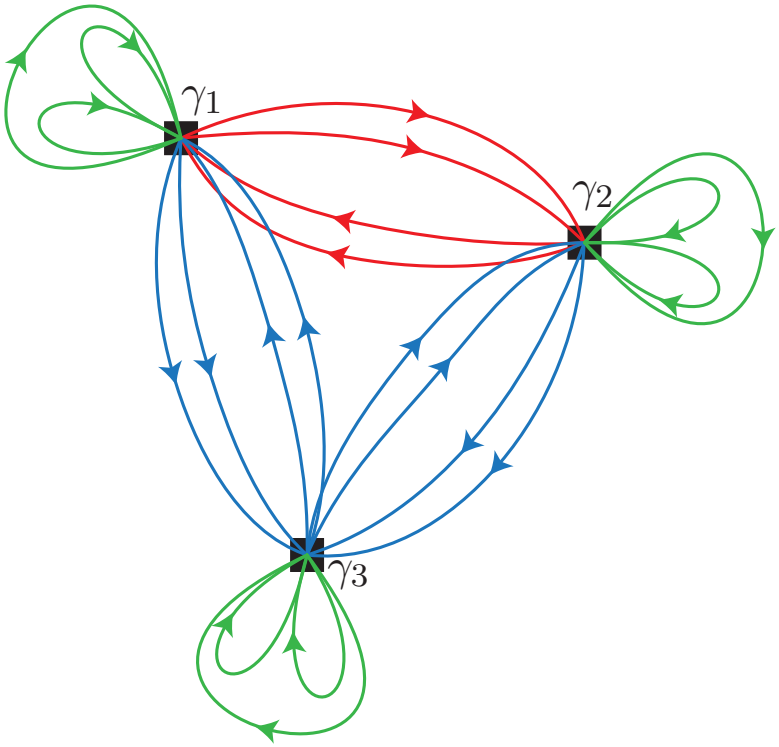


Figure 7.18: A depiction of the (infinitely many) homo- and heteroclinic orbits which connect periodic orbits of $\sigma$.

In fact there also exists an uncountable infinity of non-periodic (not even asymptotically periodic) orbits in $\Sigma$. To show this, we will show that $\Sigma$ is uncountable, i.e. that it cannot be arranged into a well-defined, countable sequence.

**Definition 7.10.** The binary expansion of a number $x \in [0,1]$ is the sequence $(b_i)_{i=1}^{\infty}$ of 1s and 0s such that

$$x = \sum_{i=1}^{\infty} \frac{b_i}{2^i}. \tag{7.48}$$

To calculate this sequence define the map $g(x) = 2x \mod 1$. Then generate the sequence $(x_i)_{i=1}^{\infty} = (x, g(x), g^2(x), \ldots)$. Finally, set each binary coefficient as

$$b_i = \begin{cases} 0 & 2x_i < 1 \\ 1 & 2x_i \geq 1 \end{cases}. \tag{7.49}$$

**Proposition 7.16.** *The set $\Sigma$ is uncountable.*

*Proof.* It is enough to show that the set $\tilde{\Sigma}$ defined by

$$\tilde{\Sigma} = \{.s_1 s_2 s_3 \ldots : \ s_i \in \{0,1\}\} \tag{7.50}$$

is uncountable. This suffices as we can injectively embed $\tilde{\Sigma}$ into $\Sigma$ via the map $.s_1 s_2 \ldots \mapsto \overline{1}.(s_1+1)(s_2+1)\ldots$.

Every binary expansion for all real numbers in the interval $[0,1]$ is contained within $\tilde{\Sigma}$, therefore the cardinality of $\tilde{\Sigma}$ is at least as large as that of $[0,1]$ which is uncountable. In fact, the binary expansion is a bijection, and these two sets have the exact same cardinality. □

*Example* 7.4 (Binary expansion). Let $x = \frac{1}{5}$, we find the iterates of $g$

$$x = \frac{1}{5} \qquad 2 \cdot \frac{1}{5} < 1 \implies x_1 = 0 \tag{7.51a}$$

$$g\left(\frac{1}{5}\right) = \frac{2}{5} \qquad 2 \cdot \frac{2}{5} < 1 \implies x_2 = 0 \tag{7.51b}$$

$$g^2\left(\frac{1}{5}\right) = g\left(\frac{2}{5}\right) = \frac{4}{5} \qquad 2 \cdot \frac{4}{5} > 1 \implies x_3 = 1 \tag{7.51c}$$

$$g^3\left(\frac{1}{5}\right) = g\left(\frac{4}{5}\right) = \frac{8}{5} \mod 1 = \frac{3}{5} \qquad 2 \cdot \frac{3}{5} > 1 \implies x_4 = 1 \tag{7.51d}$$

$$g^4\left(\frac{1}{5}\right) = g\left(\frac{3}{5}\right) = \frac{6}{5} \mod 1 = \frac{1}{5} \qquad 2 \cdot \frac{1}{5} < 1 \implies x_5 = 0. \tag{7.51e}$$

Since $g^4\left(\frac{1}{5}\right) = \frac{1}{5}$, we have reached a repeating loop (since our seed value was $\frac{1}{5}$), therefore we have that

$$\frac{1}{5} = 0 \cdot \frac{1}{2} + 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{8} + 1 \cdot \frac{1}{16} + \ldots = .\overline{0011}. \tag{7.52}$$

*Remark* 7.17. Recall that for a metric space $(X,d)$ when we write $\mathcal{B}(x,r)$ we refer to the open ball of radius $r$ around $x$, i.e.

$$\mathcal{B}(x,r) = \{y \in X : \ d(x,y) < r\}. \tag{7.53}$$

**Proposition 7.18.** *The map $\sigma$ has a dense orbit in $\Sigma$, i.e. an orbit which approaches any point in $\Sigma$ arbitrarily closely.*

**Proposition 7.19.** *The map $\sigma$ is topologically transitive, i.e. for any two open sets $A$ and $B$ in $\Sigma$, there exists an $N \in \mathbb{N}$ such that*

$$\sigma^N(A) \cap B \neq \emptyset. \tag{7.54}$$

This is also called the *mixing property*. In the case of the Bernoulli shift map, it can be concluded from the existence of a dense orbit.

**Proposition 7.20** (Sensitive dependence on initial conditions)**.** *Regardless of how close two distinct initial conditions $s$ and $s'$ are chosen, they will not remain close. There exists a $\Delta > 0$ such that for every $s \in \Sigma$ and any $\delta > 0$, there exists an $s' \in \mathcal{B}(s, \delta)$ such that for an $N$ large enough the following holds*

$$d(\sigma^N(s), \sigma^N(s')) > \Delta. \tag{7.55}$$

*This is called the* uniform instability of all orbits. *This fact is depicted in Fig. 7.19.*

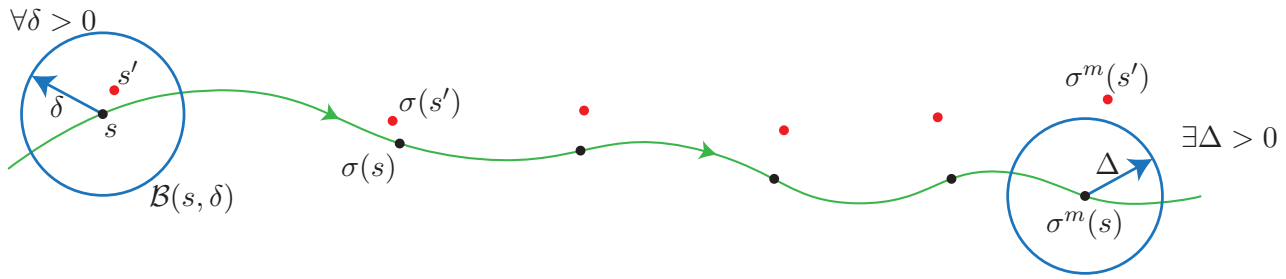


Figure 7.19: The uniform instability of orbits for the map $\sigma$.

*Remark* 7.21. This by itself implies no complexity: the linear saddle

$$\begin{cases} \dot{x} = \lambda x \\ \dot{y} = -\mu y, \end{cases} \tag{7.56}$$

also has this property.

*Remark* 7.22. In the case of this map, every $s' \in \mathcal{B}(s, \delta)$ has an $N$ large enough such that the inequality holds.

These observations lead us to the following definition of a chaotic map.

**Definition 7.11** (Chaotic map)**.** Let $M : C \to C$ be a $\mathcal{C}^0$ map on a compact metric space $C$. We call $M$ *chaotic* if

(i) $M$ is topologially transitive (the mixing property);

(ii) $M$ has sensitive dependence on initial conditions (a lack of reproducibility);

(iii) $M$ has a dense set of periodic orbits in $C$.

The Bernoulli shift map $\sigma$ is chaotic, therefore Smale's horseshoe map on the invariant set $\Lambda$ is chaotic, therefore the horseshoe map on the unit square is chaotic. Since that was a model for the Poincaré map with a homoclinic tangle, we can rigorously conclude that the forced Duffing equation is chaotic. All this followed from a transverse zero of the Melnikov function.

*Example* 7.5 (Chaos in a 1-dimensional model). Consider the logistic map for resource-limited discrete population growth dynamics

$$x_{n+1} = x_n = ax_n(1 - x_n); \quad a \in \mathbb{R}^+. \tag{7.57}$$

We will now continue with $a = 4$. The points $x = 0$ and $x = \frac{3}{4}$ are fixed points. These are each unstable as $|f'(x)| > 1$ at both of these points. This is not an ODE, and we do not assume that the model is a small perturbation of a simpler system. That is, we are not assuming persistence of known structures. The cobweb diagram is illustrated in Fig. 7.20. We can see that for our

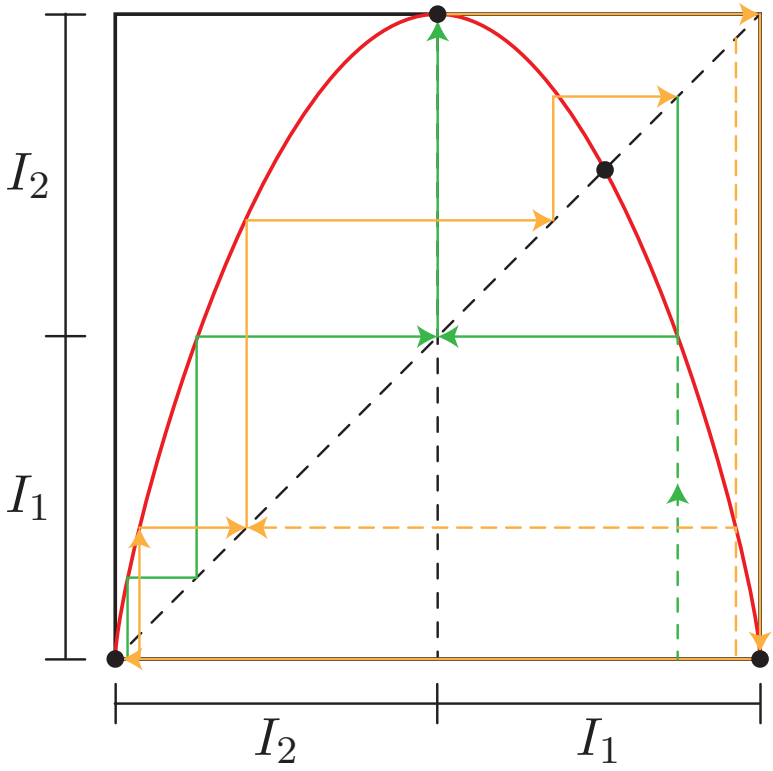


Figure 7.20: The cobweb diagram for the logisitic map with $a = 4$. The sets $I_1$ and $I_2$ are defined as $[0, \frac{1}{2}]$ and $[\frac{1}{2}, 1]$.

choice of $a$, $f$ maps $[0, 1]$ to $[0, 1]$ and is not invertible. In fact each preimage of $\frac{1}{2}$ reaches 0 in a finite number of iterations. Also note that because the map is quadratic, each point has

two preimages. Following a point in backward time is not well defined in the usual sense, but instead, we may take preimages of preimages of $\frac{1}{2}$. This gives rise to two branches of backward iterates, i.e. two chains of preimages of $\frac{1}{2}$. One of these branches (branch A) converges to 0, and the other one (branch B) converges to $\frac{3}{4}$ in backward time. Points on branch A thus form a homoclinic orbit to zero, and points on branch B form a heteroclinc orbit between $\frac{3}{4}$ and 0. Since a point may follow branch A for a finite number of backward iterations before switching to branch B and vice versa, we actually have a countably infinite number of homo- and heteroclinic orbits. From Fig. 7.20, the dynamics appears complicated, with no stable fixed points as we have seen from the linearization.

To better understand the dynamics we can encode the itinerary of a point $x \in I = [0,1]$ by if $x$ is in $I_1$ or $I_2$. We get a one-sided itinerary, because $f$ is not invertible. For example if $x \in I_{s_0}$, $f(x) \in I_{s_1}$, $f^2(x) \in I_{s_2}$ then we would encode $x$ as $\bullet s_0 s_1 s_2 \ldots$ for $s_i \in \{1,2\}$.

Note here that this symbolic encoding is not well defined for every point in $I$, as all preimages of $x = \frac{1}{2}$ have multiple encodings, since their iterates eventually reach $I_1 \cap I_2 = \frac{1}{2}$. To address this, instead we hope to show that each encoding corresponds to a unique trajectory. To this end define $\Psi : \Sigma \to I$ for $\Sigma$ the set of semi-infinite binary symbols, by

$$\Psi : \bullet s_0 s_1 s_2 \mapsto x = I_{s_0} \cap f^{-1}(I_{s_1}) \cap f^{-2}(I_{s_2}) \cap \ldots. \tag{7.58}$$

We now have to ask if this is well defined, i.e. if it yields a unique $x$ value. Note here that for any subset $J \subset I$ which is closed and connected, we have that $I_{s_i} \cap f^{-1}(J)$ is closed, nonempty, and connected. This is due to the symmetry around $x = \frac{1}{2}$ and is shown visually in Fig. 7.21.

*Remark* 7.23. For any two sets $A$ and $B$ we have that

$$f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B). \tag{7.59}$$

Further observe that for $I_{s_0 s_1} = I_{s_0} \cap f^{-1}(I_{s_1})$ is closed, nonempty, and connected for all $s_0$ and $s_1$ by the previous note. Continuing this we find that

$$I_{s_0 s_1 s_2} = I_{s_0} \cap f^{-1}(I_{s_1}) \cap f^{-2}(I_{s_2}) = I_{s_0} \cap f^{-1}\left(I_{s_1} \cap f^{-1}(I_{s_2})\right) \tag{7.60}$$

is closed, nonempty, connected, and contained in $I_{s_0 s_1}$. This line of argument can be continued to show for any $n \in \mathbb{N}$ the set $I_{s_0 \ldots s_n}$ is closed, nonempty, connected, and contained in $I_{s_0 \ldots s_{n-1}}$. Thus we have a nested sequence of closed, nonempty, and connected sets. By Cantor's theorem a nonempty limit exists, which is now not a Cantor set in this case, because it is connected. It is also possible to show that $\lim_{n\to\infty} |I_{s_0 \ldots s_n}| = 0$, where $|\cdot|$ denotes the length of a interval. Now we can can conclude that the limit set contains a single point $x$, therefore as $n \to \infty$ the function $\Psi(s) = x$ is well-defined. It is also true that every element $x \in I$ can be written in this form for some $s$, i.e. $\Psi$ is a surjection.

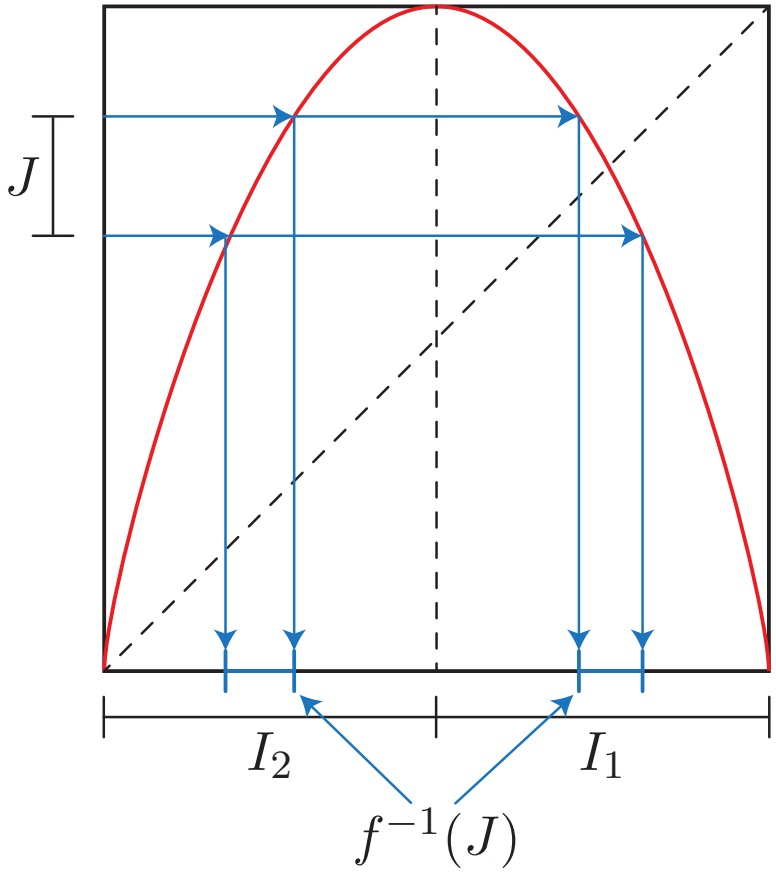


Figure 7.21: The preimage of a subset $J \subset I$, with its closed, nonempty, and connected intersections with $I_1$ and $I_2$.

Additionally with the metric

$$d(s, s') = \sum_{i=0}^{\infty} \frac{|s_i - s_i'|}{2^i} \tag{7.61}$$

the function $\Psi$ is continuous.

By definition, a single iteration of the map $f$ is equivalent to a shift to the left on the corresponding symbols in $\Sigma$, i.e.

$$f : x = .s_0 s_1 s_2 \ldots \mapsto .s_1 s_2 \ldots = f(x). \tag{7.62}$$

Therefore we have the following commuting diagram

$$\begin{array}{ccc} \Sigma & \xrightarrow{\sigma} & \Sigma \\ \downarrow{\scriptstyle \Psi} & & \downarrow{\scriptstyle \Psi} \\ I & \xrightarrow{f} & I \end{array} . \tag{7.63}$$

*Definition* 7.12 (Topological semi-conjugacy). A map $f$ is *topologically semi-conjugate* to $\sigma$ if there exists a surjection $\Psi$ such that

$$\boxed{f \circ \Psi = \Psi \circ \sigma.} \tag{7.64}$$

In comparison to the previous example, $\Psi$ is not invertible (although it is well-defined and continuous). The invertibility of $\Psi$ only fails on a countable set of points $\left\{f^{-n}\left(\frac{1}{2}\right)\right\}_{n=1}^{\infty}$. Removing this countable set of points, a homeomorphism between $\Sigma$ and $\tilde{I} = I \setminus f^{-n}(1/2)_{n=1}^{\infty}$ can be constructed. Therefore we have that the logistic map is chaotic on $\tilde{I}$.

## 7.4 Generalization of chaotic dynamics

We begin by defining a general symbol space.

**Definition 7.13.** Denote by $\Sigma^N$ the symbol space defined as the space of doubly infinite sequences of $N$ symbols (for $N \in \mathbb{N}$), i.e.

$$\boxed{\Sigma^N = \{s = \dots s_{-2}s_{-1}\bullet s_0 s_1 s_2 \dots : \ s_i \in \{0, 1, \dots, N-1\}\}.} \tag{7.65}$$

We may constrain this symbol space by constraining which sequences are admissible. This can be done by using a transition matrix $A \in \mathbb{R}^{N\times N}$ with entries $A_{ij} \in \{0, 1\}$. Then the constrained symbol space is given by

$$\Sigma_A^N = \left\{s \in \Sigma^N : \ A_{s_i s_{i+1}} \neq 0, \forall i\right\}. \tag{7.66}$$

In the constrained space, consecutive symbols $s_i$ and $s_{i+1}$ must correspond to a 1 in the transition matrix $A$. In other words if the symbol $s_i = k$ then the symbol $s_{i+1}$ must be equal to $j$ such that $A_{kj} = 1$.

*Example* 7.6 (Constrained symbol space). To demonstrate how a transition matrix can constrain the symbol space, lets start by examining $\Sigma_A^2$ for

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{7.67}$$

Be examining the entries of the transition matrix $A$ we find which strings are admissible

$$\begin{aligned} A_{11} = 1 &\implies \dots 00 \dots \quad \text{admissible} \\ A_{12} = 0 &\implies \dots 01 \dots \quad \text{inadmissible} \\ A_{21} = 0 &\implies \dots 10 \dots \quad \text{inadmissible} \\ A_{22} = 1 &\implies \dots 11 \dots \quad \text{admissible.} \end{aligned}$$

Therefore the set $\Sigma_A^2$ is comprised only of two elements $\{\overline{0}\bullet\overline{0}, \overline{1}\bullet\overline{1}\}$. As this space is very simple and only contains fixed points of $\sigma$, no chaos can occur.

**Definition 7.14.** The map

$$\sigma : \Sigma_A^N \to \Sigma_A^N; \quad \dots s_{-1}\bullet s_0 \dots \mapsto \dots s_{-1}s_0\bullet s_1 s_2 \dots, \tag{7.68}$$

is called a *Bernoulli subshift of finite type* (i.e. $N < \infty$) with transition matrix $A$.

Before more examination of chaos on such constrained spaces, we need to recall a definition from linear algebra.

**Definition 7.15.** A matrix $A$ is called *irreducible* if there exists a $k \in \mathbb{N}$ such that $\left(A^k\right)_{ij} \neq 0$ for every $i, j \in \{1, \ldots, N\}$.

**Theorem 7.24.** *Assume that $A$ is irreducible, then*

*(i) $\Sigma_A^N$ is a Cantor set; and*

*(ii) $\sigma : \Sigma_A^N \to \Sigma_A^N$ is a chaotic map.*

Note that in the previous example with $A = I_2$, the transition is not irreducible, which is why chaos is not guaranteed.

The role of irreducibility is critical for the theorem to work. Say for instance that the required $k$ such that each entry of $A^k$ is nonzero is $k = 3$. Then we know that $(A^3)_{ij} = \sum_{l,m=1}^{N} a_{il}a_{lm}a_{mj} \neq 0$ for any pair $i, j$. Therefore there exist $l, m$ such that $a_{il}$, $a_{lm}$, and $a_{mj}$ which are not equal to 0. Hence there exists an admissible ultimate transition from state $i$ to state $j$ for all pairs $i, j$, if the symbol sequence is viewed as an itinerary. Furthermore, different irreducible transition matrices generate different types of chaos.

*Example* 7.7 (Chaos in the Newton-Raphson iteration). Newton-Raphson iteration is a process to find the roots (zeros) of a nonlinear function $f$. The iteration is as follows

(i) Start with the seed point $x_0$;

(ii) Define $x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} = N(x_0)$;

(iii) Repeat (ii) as $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$, until convergence, i. e. until $x_{n+1} = x_n$.

In practice this approach is often truncated when $x_{n+1} \approx x_n$. The roots of $f$ coincide with the fixed points of the 1-dimensional dynamical system $x_{n+1} = N(x_n)$. If such a fixed point $\overline{x}$ is stable and $x_0$ is in its domain of attraction, then $\lim_{i\to\infty} N^i(x_0) = \overline{x}$. The iteration process is sketched in Fig. 7.22.

Note that if $N(\overline{x}) = \overline{x}$ then we have

$$N'(\overline{x}) = 1 - \frac{f'(\overline{x})^2 - f(\overline{x})f''(x_0)}{f'(\overline{x})^2} = 0. \tag{7.69}$$

The latter equality is due to $f(\overline{x}) = 0$. Stability of $\overline{x_0}$ clearly follows, because $|N'(\overline{x_0})| < 1$. Additionally, since $N'(\overline{x_0}) = 0$, we call the fixed point *super-stable*.

Specifically, once we are near the optimum $\overline{x}$, we have

$$N(x_n) = \underbrace{N(\overline{x})}_{=\overline{x}} + \underbrace{N'(\overline{x})}_{=0}(x_n - \overline{x}) + \frac{1}{2}N''(\overline{x})(x_n - \overline{x})^2 + \mathcal{O}\left(|x_n - \overline{x}|^3\right). \tag{7.70}$$

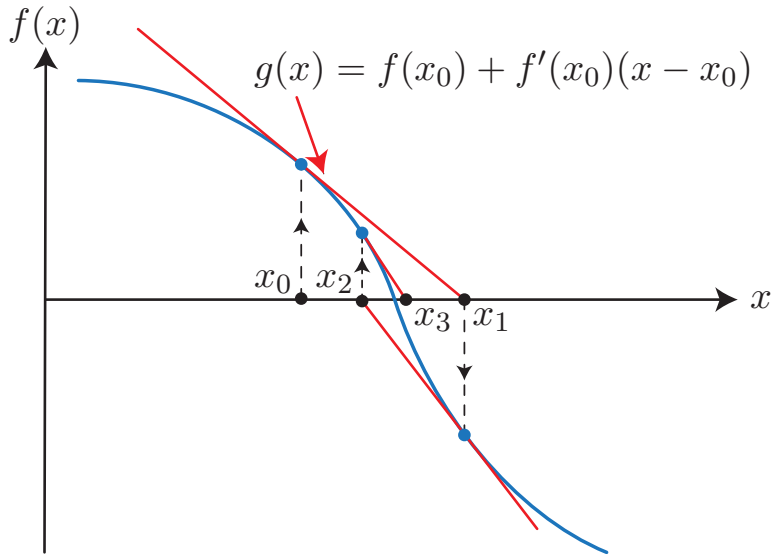


Figure 7.22: An illustration of the iteration of the Newton-Raphson process on the function $f$ (blue). The red designates the function $g$, when given the point $x_0$ the next iterate $x_1$ fufills the property $g(x_1) = 0$.

This implies we have fast quadratic convergence, as

$$|N(x_n) - N(\overline{x})| \approx \frac{1}{2}|N''(\overline{x})||x_n - \overline{x}|^2. \tag{7.71}$$

Despite the fast convergence, erratic iterations may result if the initial point is not close enough to the root [Saari and Urenko, 1984]. In particular, consider a 4th order polynomial

$$f(x) = (x - x_0)(x - x_1)(x - x_2)(x - x_3); \quad x_0 < x_1 < x_2 < x_3. \tag{7.72}$$

This can be observed to lead to chaotic Newton-Raphson iterations. The proof will be sketched for this chaotic behaviour.

Retain the 4th order polynomial $f$ from above. We then have $N(x_i) = x_i$ for $i = 0, 1, 2, 3$ as these are roots of $f$. Note then that we already know what the shapes of $f$ and $f'$ are, and these are illustrated in Fig. 7.23.

The points $a_i$ denote the minima and maxima of $f(x)$ and $b_i$ denote the minima and maxima of $f'(x)$ (See Fig. 7.23). As $f'(a_i) = 0$ we also have that $|N(a_i)| = \infty$, also $f''(b_i) = 0$ implies that $N'(b_i) = 0$. Thus via our knowledge of the shape of $f$, we have been able to derive information about $N$. If we assume that $b_1 \in (a_0, x_1)$ and $b_2 \in (x_2, a_2)$, then we get the graph of $N$ as shown in Fig. 7.24.

Next, we seek non-convergent iterations in $[a_0, a_1] \cup [a_1, a_2]$. We begin by finding the preimages of $a_0$ (called $y_{0,i}$) and $a_2$ (called $y_{2,i}$) this can be done geometrically by starting from the

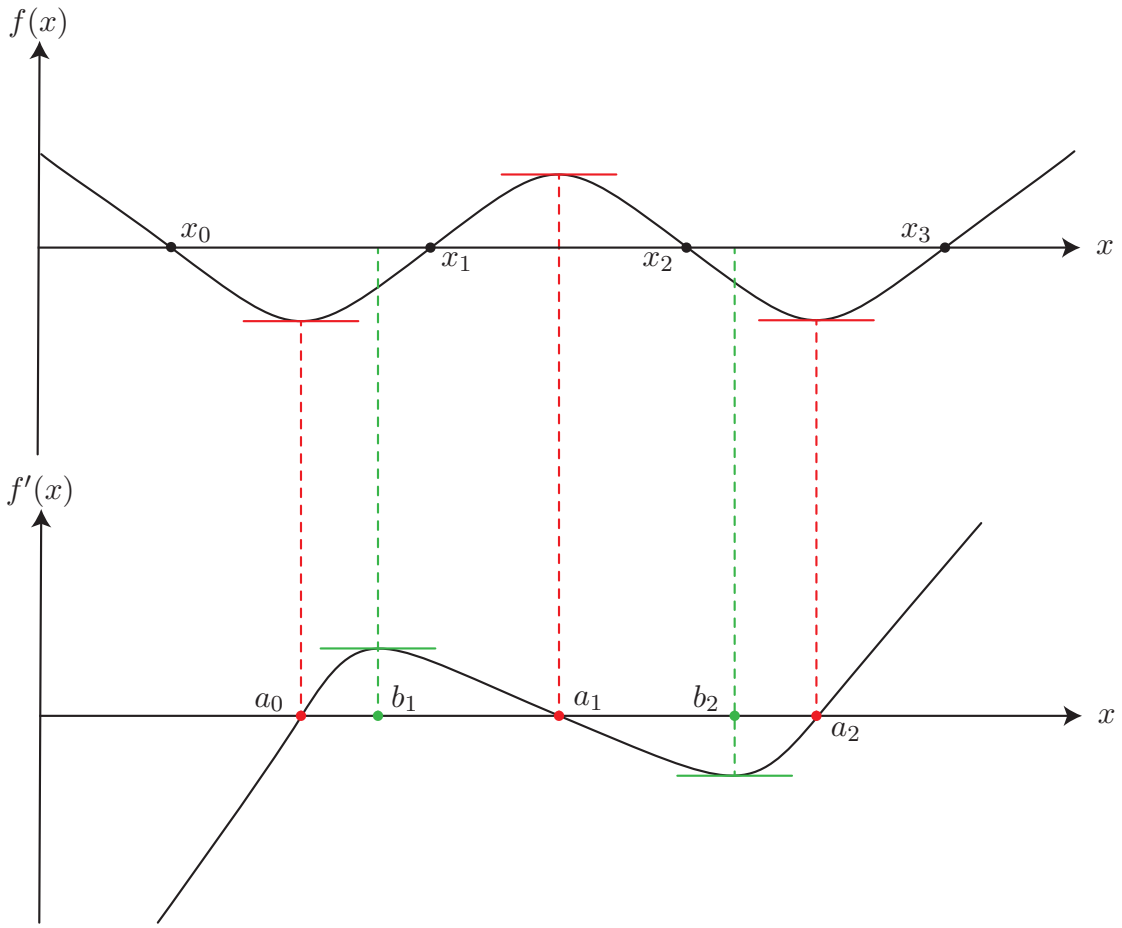


Figure 7.23: The shape of $f$ (above) and $f'$ (below). The points $a_i$ are roots of $f'$ and the points $b_i$ are roots of $f''$.

point $a_k$, going down to the diagonal, and then searching horizontally for where the graph of $N$ crosses the level $a_k$; going down from this point to the $x$-axis yields a preimage. This is shown in Fig. 7.25. Using these points define the following sets

$$I_1 = [y_{2,1}, z_1^+]; \quad I_2 = [z_1^-, y_{0,1}]; \quad I_3 = [y_{2,2}, z_2^+]; \quad I_4 = [z_2^-, y_{0,2}]. \tag{7.73}$$

These sets are illustrated in Fig. 7.25 as well.

Noting that $b_2 > N(b_2)$ and $b_1 < N(b_1)$ observe that the following hold

$$N(I_1) = [x_1, a_2] \supset I_2 \cup I_3 \cup I_4; \quad N(I_3) = [N(b_2), a_2] \supset [b_2, a_2] \supset I_4; \tag{7.74a}$$

$$N(I_2) = [a_0, N(b_1)] \supset [a_0, b_1] \supset I_1; \quad N(I_4) = [a_0, x_2] \supset I_1 \cup I_2 \cup I_3. \tag{7.74b}$$

Therefore we have that the possible transitions from $I_i$ to $I_j$ are described by the transition

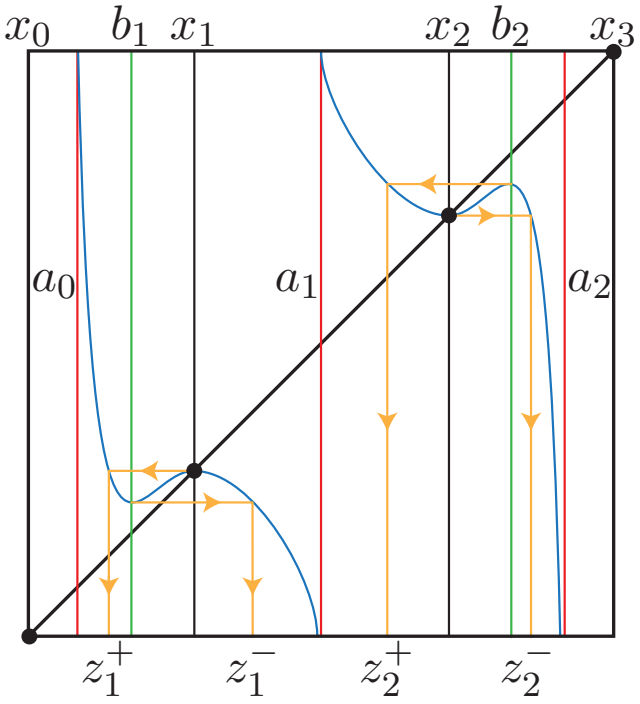


Figure 7.24: The graph of $N$, the points $z_i^{\pm}$ are lower bounds on the domain of attraction as they surely converge to the fixed points. The are chosen by starting at a local extrema and working backwards.

matrix

$$A = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix}. \tag{7.75}$$

This transition matrix is in fact irreducible with $k = 3$

$$A^2 = \begin{pmatrix} 2 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 2 \end{pmatrix}; \quad A^3 = \begin{pmatrix} 2 & 3 & 3 & 3 \\ 2 & 1 & 1 & 1 \\ 1 & 1 & 1 & 2 \\ 3 & 3 & 3 & 2 \end{pmatrix}. \tag{7.76}$$

We are not done yet, as we still need to establish a unique encoding. Note that if $a_{ij}^3 \neq 0$

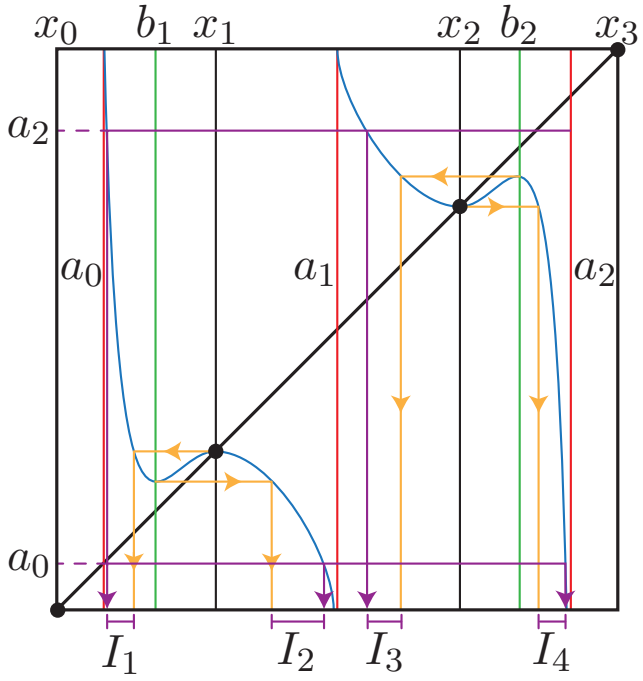


Figure 7.25: The preimages of the values $a_0$ and $a_2$, along with the sets $I_j$ for $j = 1, 2, 3, 4$.

then the set $I_i \cap f^{-1}(I_j) \neq \emptyset$ is closed for $i = 1, 2, 3, 4$. In light of this, with $\tilde{N} = N^3$, define

$$I_{s_0 s_1 \ldots s_n} = I_{s_0} \cap \tilde{N}^{-1}(I_{s_1}) \cap \tilde{N}^{-2}(I_{s_2}) \cap \ldots \cap \tilde{N}^{-n}(I_{s_n}) \tag{7.77a}$$

$$= I_{s_0} \cap \tilde{N}^{-1}\left(I_{s_1} \cap \tilde{N}^{-1}\left(I_{s_2} \cap \ldots \tilde{N}^{-1}\underbrace{\left(I_{s_{n-1}} \cap \tilde{N}^{-1}(I_{s_n})\right)}_{\neq\emptyset,\text{ closed}}\right)\right). \tag{7.77b}$$

The condition $I_{s_{n-1}} \cap \tilde{N}^{-1}(I_{s_n}) \neq \emptyset$ and is closed is true as each entry $a^3_{ij} \neq 0$. Therefore we have a nested sequence of nonempty, closed sets and by Cantor's Theorem their intersection yields a nonempty, closed set. By our previous result we have that $\Lambda = \bigcup_{s \in \Sigma_A^N} I_s$ is a Cantor set. In this case $N = 4$ and $A$ is as above. We then have the encoder function $\phi : \Lambda \to \Sigma_A^4$ and the chaotic function $\sigma : \Sigma_A^4 \to \Sigma_A^4$. These yield the commuting diagram

$$\begin{array}{ccc} \Sigma_A^4 & \xrightarrow{\sigma} & \Sigma_A^4 \\ \phi\uparrow & & \phi\uparrow \\ \Lambda & \xrightarrow{N} & \Lambda \end{array} . \tag{7.78}$$

Thus we have that $N = \phi^{-1} \circ \sigma \circ \phi$, which implies that $N$ is a chaotic map. Hence, the

Newton-Raphson iteration has an invariant Cantor set, on which the iteration is equivalent to a Bernoulli subshift of finite type on 4 symbols.

# Chapter 8

# Introduction to Hamiltonian Systems

A canonical Hamiltonian system is given by two state variables $p, q \in \mathbb{R}^n$ (i.e. the entire system is even dimensional), a *Hamiltonian* function $H(p,q,t) \in \mathcal{C}^1(\mathbb{R}^{2n} \times \mathbb{R})$ which maps $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$. In general $q$ denotes the generalized coordinates (position) and $p$ the generalized momenta. The dynamics on the system is given by

$$\begin{cases} \dot{q} = \dfrac{\partial H(q,p,t)}{\partial p} \\ \dot{p} = -\dfrac{\partial H(q,p,t)}{\partial q}. \end{cases} \tag{8.1}$$

This is represented in shorthand notation by using the identity matrix, denoted by $I_n \in \mathbb{R}^{n \times n}$ and the symplectic matrix

$$J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}. \tag{8.2}$$

Denote the differential with respect to the state variables $x = \begin{pmatrix} p \\ q \end{pmatrix} \in \mathbb{R}^{2n}$ by

$$D_x = D = \begin{pmatrix} \frac{\partial}{\partial q} \\ \frac{\partial}{\partial p} \end{pmatrix}. \tag{8.3}$$

Then the shorthand notation for the system is

$$\boxed{\dot{x} = JD_xH(x,t).} \tag{8.4}$$

*Example* 8.1 (Holonomic mechanical systems are Hamiltonian). Assume that all external forces are derived from a potential, i.e. $F(q,t) = -\frac{\partial V(q,t)}{\partial q}$ where $q$ denotes the position. Further

assume that all constraints are geometric, i.e. $f(q,t) = 0$. We have $n$ degrees of freedom as $q \in \mathbb{R}^n$ and $\dot{q} \in \mathbb{R}^n$. The kinetic energy is given by $T(q,\dot{q})$ and the potential energy by $V(q,t)$. We then use the principle of least action (also referred to as Hamilton's principle) to derive the equations of motion. This can be shown to be equivalent to Newton's second law. This principle states that for the *Lagrangian* $L(q,\dot{q},t) = T(q,\dot{q}) - V(q,t)$ we have that the variation of the action $S$ is 0

$$\boxed{\delta S = \delta \int_{t_0}^{t_1} L(q(s),\dot{q}(s),s)ds = 0.} \tag{8.5}$$

This implies the *Euler-Lagrange equations*

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0. \tag{8.6}$$

Thus we have $n$, second-order ODEs.

*Remark* 8.1. In (8.5) the variation $\delta S$ being equal to 0 stems from the principle of least action, as the action $\int L(q(s),\dot{q}(s),s)ds$ is then a minimum. The variational of a function $\delta F$ (here $F[\gamma] = \int \gamma ds$) of a path (in our case $\gamma = (q(s),\dot{q}(s),s)$) corresponds to seeing how infinitesimal changes to the path change the output of the functional. These different, admissible paths are depicted in Fig. 8.1. Hamilton's principle states that among these different paths the one that will physically take place is the one along which the action $S$ is minimal.

Now assume that $L$ is a convex function of $\dot{q}$. Using this the *Legendre transfomation* of $L$ can be performed, which is essentially just a different way to look at the graph of $L(q,\cdot,t)$. This transformation gives us

$$\boxed{H\left(q, \frac{\partial L}{\partial \dot{q}}, t\right) = \frac{\partial L}{\partial \dot{q}}\dot{q} - L(q,\dot{q},t).} \tag{8.7}$$

Th geometric interpretation of this is shown in Fig. 8.2. At each velocity $\dot{q}$ the tangent line of $L(q,\cdot,t)$ is taken and shifted such that the $y$-intercept is at 0. Then $H$ is equal to the difference between this shifted line and the function $L(q,\dot{q},t)$.

As $L$ is convex, if we define the generalized momenta $p = \frac{\partial L}{\partial \dot{q}}$. There then exists a unique $\dot{q}$ such that $\dot{q} = F(q,p,t)$, in fact the relationship between $\dot{q}$ and $p$ here is bijective. We then have

$$H(q,p,t) = pF(q,p,t) - L(q,F(q,p,t),t) \tag{8.8}$$

the Hamiltonian associated with $L$ (or the Legendre transform of $L$).

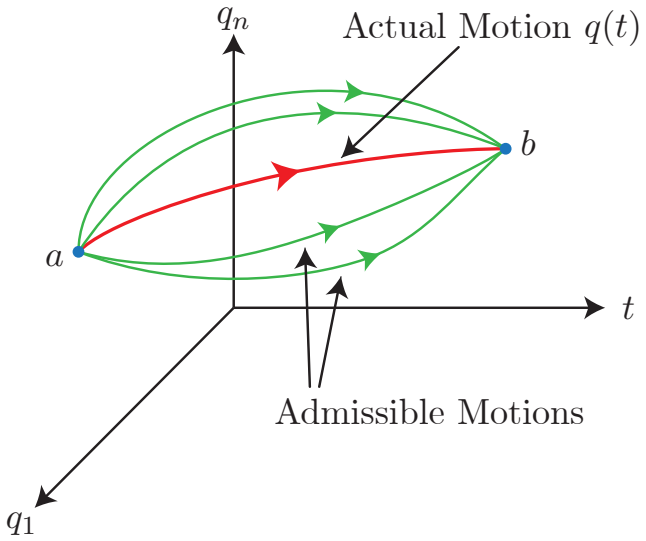


Figure 8.1: An illustration of various different paths (denoted here by $q(t)$). The red demarcates the actual motion which occurs in the system, whereas the green paths are each admissible motions which are consistent with the constraints.

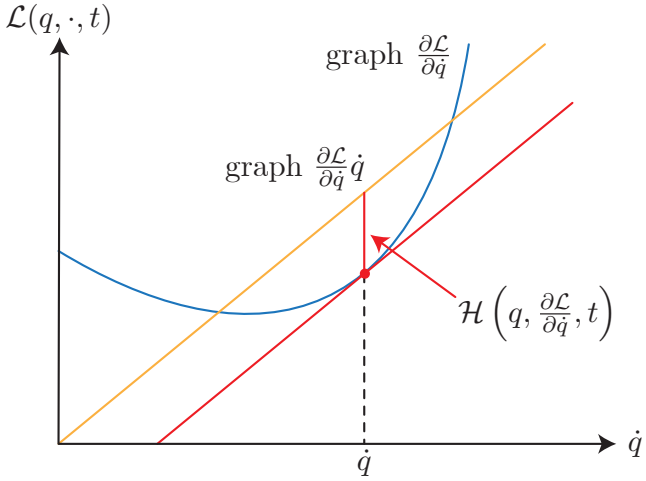


Figure 8.2: The Legendre transformation. The thin red linear function denotes the tangent of $L(q,\cdot,t)$ taken at $\dot{q}$, the yellow linear function is then the shifted version with a 0 $y$-intercept, and the thick red vertical line is the Legendre transformation $H$ measuring the difference between this shifted line and $L(q,\cdot,t)$.

At this point note that the total differential of $H(q,p,t)$ is

$$dH = \frac{\partial H}{\partial q}dq + \frac{\partial H}{\partial p}dp + \frac{\partial H}{\partial t}dt \tag{8.9a}$$

$$= d(pF) - \underbrace{\frac{\partial L}{\partial q}}_{=\dot{p}} dq - \underbrace{\frac{\partial L}{\partial \dot{q}}}_{=p} - \frac{\partial L}{\partial t}dt \tag{8.9b}$$

$$= \underbrace{F}_{=\dot{q}} dp + pdF - \dot{p}dq - pdF - \frac{\partial L}{\partial t}dt \tag{8.9c}$$

$$= \dot{q}dp - \dot{p}dq - \frac{\partial L}{\partial t}dt. \tag{8.9d}$$

Comparing coefficients in the first and last equations then yields

$$\begin{cases} \dot{q} = \dfrac{\partial H}{\partial p} \\ \dot{p} = -\dfrac{\partial H}{\partial q} \end{cases}. \tag{8.10}$$

This first-order ODE is called the *Hamiltonian equations of motion*. The physical meaning of the Hamiltonian can be seen by substituting $L = T(q, \dot{q}) - V(q, t)$ into the definition 8.7. The kinetic energy only depends on $\dot{q}$ through the kinetic energy,

$$\dot{q}\frac{\partial L}{\partial \dot{q}} = \dot{q}\frac{\partial T}{\partial \dot{q}}. \tag{8.11}$$

Since the kinetic energy is quadratic in the velocity, this simplifies to

$$\dot{q}\frac{\partial T}{\partial \dot{q}} = 2T(q, \dot{q}). \tag{8.12}$$

We then have

$$H = 2T - (T - V) = T(q, p) + V(q, t), \tag{8.13}$$

which is the full mechanical energy. Also note that we have $\frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}$.

We will now apply this result to a well known system.

*Example* 8.2 (The Hamiltonian of the pendulum). Recall that for the pendulum the position is given by $q = \varphi \in S^1$, with a string of length $\ell$ the vertical deflection of the pendulum is $\ell(1 - \cos(\varphi)$. The string is assumed massless and the mass at the end of the string is given by $m$.

We have the kinetic and potential energies

$$T = T(\dot{q}) = \frac{1}{2}m(\ell\dot{\varphi})^2 = \frac{1}{2}m\ell^2\dot{\varphi}^2; \quad V = V(q) = mg\ell(1 - \cos(\varphi)). \tag{8.14}$$

Therefore the Lagrangian is given by

$$L = T - V = \frac{1}{2}m\ell^2\dot{\varphi}^2 - mg\ell(1 - \cos(\varphi)). \tag{8.15}$$

To find the generalized momentum, we calculate the partial derivative of the Lagrangian with respect to $\dot{q}$, i.e.

$$p = \frac{\partial L}{\partial \dot{q}} = \frac{\partial L}{\partial \dot{\varphi}} = m\ell^2\dot{\varphi} \implies \dot{\varphi} = \frac{p}{m\ell^2}. \tag{8.16}$$

Now it is possible to calculate the Hamiltonian

$$H = p\frac{p}{m\ell^2} - \frac{1}{2}m\ell^2\frac{p^2}{(m\ell^2)^2} + mg\ell(1-\cos(\varphi)) = \underbrace{\frac{1}{2}\frac{p^2}{m\ell^2}}_{T} + \underbrace{mg\ell(1-\cos(\varphi))}_{V}. \tag{8.17}$$

Hence for this Hamiltonian system we have the Hamiltonian equations of motion

$$\begin{cases} \dot{\varphi} = \dfrac{\partial H}{\partial p} = \dfrac{p}{m\ell^2} \\ \dot{p} = -\dfrac{\partial H}{\partial q} = -mg\ell\sin(\varphi). \end{cases} \tag{8.18}$$

*Example* 8.3 (2-dimensional incompressible fluids are Hamiltonian). We begin on a domain $\mathcal{D}$ contained on the plane which is simply connected, i.e. it has no holes. For each position $x \in \mathcal{D}$ we have the velocity field

$$V(x,t) = \begin{pmatrix} u(x,y,t) \\ v(x,y,t) \end{pmatrix}, \tag{8.19}$$

i.e. for each position we have

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} u \\ v \end{pmatrix}. \tag{8.20}$$

This is depicted in 8.3.

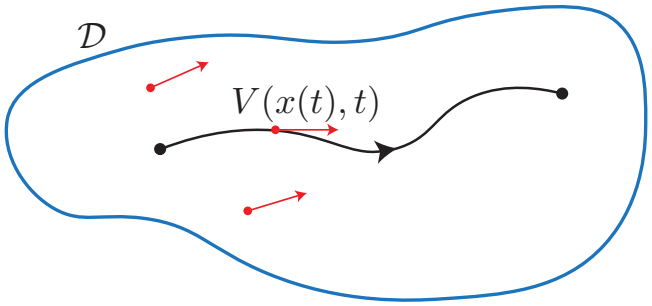


Figure 8.3: The simply connected domain $\mathcal{D}$ with the velocity field $V$.

The incompressibility condition is given by

$$\nabla \cdot V = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0 \quad \forall x \in \mathcal{D} \subset \mathbb{R}^2. \tag{8.21}$$

Now rewrite (8.21) in terms of a three dimensional vector field as

$$\nabla \times \begin{pmatrix} -v \\ u \\ 0 \end{pmatrix} = 0 \Leftrightarrow \nabla \cdot V = 0. \tag{8.22}$$

The third component is calculated as $\frac{\partial u}{\partial x} - \frac{\partial(-v)}{\partial y} = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0$. On $\mathcal{D} \times \mathbb{R}$ (which is simply connected in $\mathbb{R}^3$) the vector field $\begin{pmatrix} -v \\ u \\ 0 \end{pmatrix}$ has zero curl. Therefore there exists a potential of this vector field $\Psi : \mathcal{D} \times \mathbb{R} \to \mathbb{R}$ such that

$$\begin{pmatrix} -v \\ u \\ 0 \end{pmatrix} = \nabla\Psi = \begin{pmatrix} \partial_x \Psi \\ \partial_y \Psi \\ \partial_z \Psi \end{pmatrix}. \tag{8.23}$$

Because of the third equation $\Psi$ has no z dependence and we can write our potential as $\Psi = \Psi(x, y, t)$. Therefore we find

$$\begin{cases} -\dot{y} = \dfrac{\partial \Psi}{\partial x} \\ \dot{x} = \dfrac{\partial \Psi}{\partial y} \end{cases} \implies \begin{cases} \dot{x} = \dfrac{\partial \Psi(x,y,t)}{\partial y} \\ \dot{y} = -\dfrac{\partial \Psi(x,y,t)}{\partial x}. \end{cases} \tag{8.24}$$

Thus we have shown that particle motion is Hamiltonian.

*Example* 8.4 (Planar, autonomous dynamical systems with a conserved quantity are Hamiltonian). We begin with the dynamical system which has a first integral $U$ (which is not constant), i.e.

$$\dot{x} = f(x); \quad x \in \mathbb{R}^2; \quad \exists U : \mathbb{R}^2 \to \mathbb{R} : \ \frac{dU(x(t))}{dt} = 0. \tag{8.25}$$

Note that this first integral property has further implications

$$\frac{dU(x(t))}{dt} = DU(x(t)) \cdot \dot{x}(t) = DU \cdot f|_{x(t)} = 0 \implies DU \perp f. \tag{8.26}$$

Therefore, we can take the orthogonal complement of $DU$ and scale this to be equal to $f$, i.e.

$$f(x) = p(x) J DU(x) \implies \dot{x} = p(x) J DU(x). \tag{8.27}$$

Such a system is called a *generalized Hamiltonian system* with the scalar function $p(x)$. Here the symplectic matrix in 2 dimensions is given by the $90^\circ$ rotation

$$J = \begin{pmatrix} 0 & I_1 \\ -I_1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{8.28}$$

If we assume $p(x) \neq 0$ for all $x \in \mathcal{D} \subset \mathbb{R}^2$ then we may define the trajectory-dependent rescaling of time

$$\tau = \int_0^t p(x(s))ds. \tag{8.29}$$

Then, define the notation for the derivative with respect to $\tau$ as follows

$$\frac{d(\ )}{dt} = \frac{d(\ )}{d\tau} \underbrace{\frac{d\tau}{dt}}_{p(x(t))} = (\ )'p. \tag{8.30}$$

Hence, using this time-transformation we have

$$\dot{x} = pJDU \implies x'p = pJDU \implies \boxed{x' = JDU(x).} \tag{8.31}$$

Therefore, we have a canonical Hamiltonian system. In fact, all planar, autonomous systems with a nontrivial first integral are Hamiltonian systems after a possible reparameterization of time.

*Remark* 8.2. The fact that the first integral was not trivial (i.e. non-constant) is critical, as otherwise $DU$ would be constantly 0, as we must be able to phrase the right hand side $f$ in terms of $DU$. Only in the case $f = 0$ would this then be possible.

*Example* 8.5 (Lotka-Volterra model (1925) for predator-prey interactions). For more information about this model see [Lotka, 1925, Volterra, 1926]. Let the size of the prey population be represented by $h(t)$ and the size of the predator population be denoted by $p(t)$. Each of these populations must be at least 0. Now for strictly positive constants $a_1$, $a_2$, $b$, and $c$ we have the following dynamics

$$\begin{cases} \dot{h} = a_1 h(1 - bp) \\ \dot{p} = -a_2 p(1 - ch). \end{cases} \tag{8.32}$$

The phase portrait of this system is depicted in Fig. 8.4.

This system does in fact have a first integral, and thus it is Hamiltonian.

Now that we have seen multiple examples of Hamiltonian systems, we would like to better understand what properties they all share.

## 8.1 Properties of Hamiltonian systems

The setup going forward will be as follows for $n \geq 1$

$$\dot{x} = JDH(x,t); \quad x = \begin{pmatrix} q \\ p \end{pmatrix} \in \mathbb{R}^{2n}; \quad J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}. \tag{8.33}$$

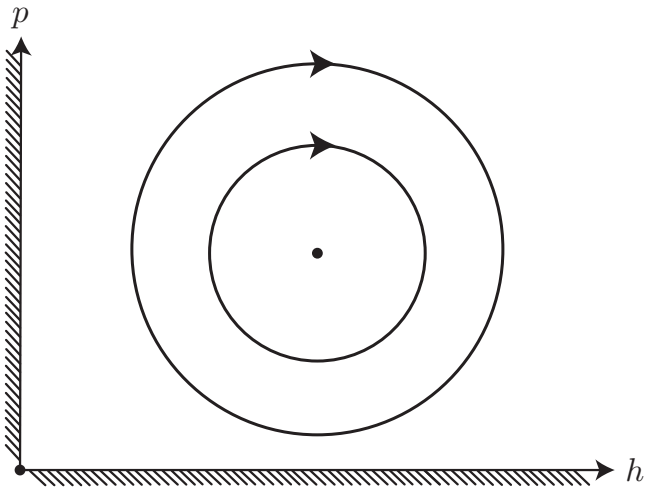


Figure 8.4: The phase portrait of the Lotka-Volterra dynamics, which has been used to explain the observed cycles in fish populations.

**Proposition 8.3** (Conservation of the Hamiltonian (energy))**.** *If* $\frac{\partial H}{\partial t} = 0$ *(i.e. H is only a function of q and p) then* $H(x(t))$ *is constant on all trajectories. Therefore* $H(x)$ *is a first integral.*

*Proof.* Use $J = -J^T$ to calculate

$$\frac{dH(x(t))}{dt} = DH(x(t)) \cdot JDH(x(t)) = \langle DH, JDH \rangle|_{x=x(t)} = 0. \tag{8.34}$$

□

The *energy surface* for some constant $E_0 = \{x \in \mathbb{R}^{2n} : \ H(x) = H_0\}$, for some constant $H_0$, is an invariant surface for the system. By nature of this $JDH$ must be tangent to $E_0$. Assuming that $E_0$ contains no fixed points of (8.33) is equivalent to $DH(x_0) \neq 0$ for all $x_0$ in $E_0$. Hence, by the Implicit Function Theorem, the solution of $H(x) - H_0 = 0$ can be smoothly continued from $x_0$. Indeed, we have

$$DH(x_0) \neq 0 \in \mathbb{R}^{2n} \implies \exists i : \ \left.\frac{\partial H}{\partial x^i}\right|_{x=x_0} \neq 0. \tag{8.35}$$

Then we have $x^i = F\left(x^1, \ldots, x^{i-1}, x^{i+1}, \ldots, x^{2n}\right)$ for $F \in \mathcal{C}^1$, because $H$ is smooth. This implies that $E_0$ is a smooth $2n-1$-dimensional graph (i.e. a smooth hypersurface or codimension-one surface).

*Example* 8.6 (Two degree of freedom nonlinear oscillator)*.* Between two fixed sides, three nonlinear springs hold two masses. Each spring has a potential given by $V_i$ for $i = 1, 2, 3$ and

the masses are given by $m_1$ and $m_2$. The displacement of each mass is given by $q_1$ and $q_2$ respectively. The potentials have global minima at $x = (q_1, q_2)^T = 0$, and are otherwise strictly monotone. The system and example potential are illustrated in Fig. 8.5.

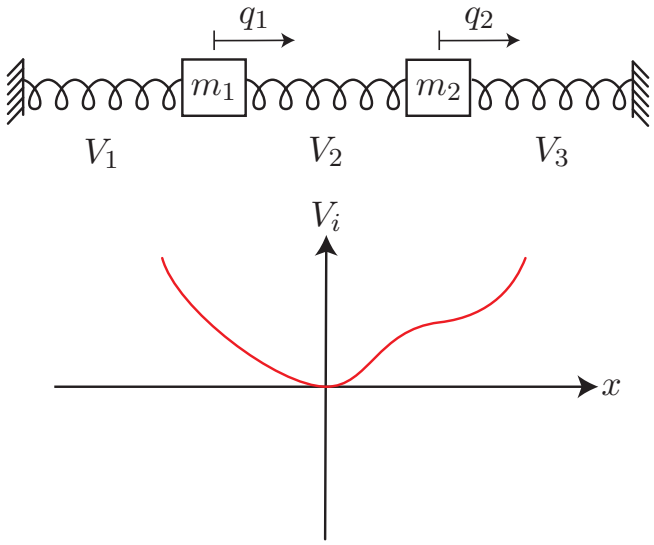


Figure 8.5: The two degree of freedom nonlinear oscillator and an example of a potential $V_i$.

For this system the specific conditions on the potentials are

$$V_i(0) = 0; \quad V_i'(0) = 0; \quad V_i''(0) > 0; \quad \forall x \neq 0 : \; V_i''(x) \neq 0. \tag{8.36}$$

The Lagrangian is then

$$L = T - V = \frac{1}{2} m_1 \dot{q}_1^2 + \frac{1}{2} m_2 \dot{q}_2^2 - \left[ V_1(q_1) + V_2(q_2 - q_1) + V_3(-q_2) \right]. \tag{8.37}$$

From the Lagrangian the generalized momenta can be derived

$$p_i = \frac{\partial L}{\partial \dot{q}_i} = m_i \dot{q}_i \implies \dot{q}_i = \frac{p_i}{m_i} \quad i = 1, 2. \tag{8.38}$$

And with this the Hamiltonian can be calculated as

$$H = T + V = \frac{1}{2} \frac{p_1^2}{m_1} + \frac{1}{2} \frac{p_2^2}{m_2} + V_1(q_1) + V_2(q_2 - q_1) + V_3(-q_2). \tag{8.39}$$

Now recall that $\dot{x} = JDH(x)$, therefore the fixed point $x_0 = (q^0, p^0)^T$ can be found by identifying the roots of $DH$

$$DH(x_0) = 0 \Leftrightarrow \begin{pmatrix} V_1'(q_1^0) - V_2'(q_2^0 - q_1^0) \\ V_2'(q_2^0 - q_1^0) - V_3'(-q_2^0) \\ \frac{p_1^0}{m_1} \\ \frac{p_2^0}{m_2} \end{pmatrix}. \tag{8.40}$$

For instance the origin is a fixed point, i.e.

$$x_0 = \begin{pmatrix} q^0 \\ p^0 \end{pmatrix} = 0 \in \mathbb{R}^4. \tag{8.41}$$

Hence we have

$$p_1^0 = p_2^0 = 0; \quad V_1'(q_1^0) = V_2^1(q_2^0 - q_1^0); \quad V_2'(q_2^0 - q_1^0) = V_3'(-q_2^0). \tag{8.42}$$

Assume that $q_1^0 > 0$ solves (8.42). Then we would have the following

$$V_1'(q_1^0) = V_2'(q_2^0 - q_1^0) > 0 \implies q_2^0 > q_1^0 > 0, \tag{8.43}$$

due to the conditions set out in (8.36). However we also have that

$$V_3'(-q_2^0) = V_2'(q_2^0 - q_1^0) > 0 \implies -q_2^0 > 0. \tag{8.44}$$

These two inequalities are clearly in contradiction. A similar argument shows that $q_1^0 < 0$ also leads to a contradiction. Therefore we find that at all equilibria $q_1^0 = 0$ and by symmetry $q_2^0 = 0$ as well. Thus the only fixed point is given by $q_1^0 = q_2^0 = 0$ and $p_1^0 = p_2^0 = 0$. We can conclude that all nonzero energy surfaces are codimension-one (3-dimensional) invariant manifolds (i.e. smooth surfaces).

Now it is possible to analyze the stability of this fixed point. To use Lyapunov's direct method, the Hamiltonian is a good candidate for a Lyapunov function, because $H(0) = 0$, $DH(0) = 0 \in \mathbb{R}^4$, and $\frac{dH(x(t))}{dt} = 0$. We still need to find out if $H$ is cup shaped. First calculate the Hessian

$$D^2H = \left(\begin{array}{c|cc} D^2V & 0_{2\times 2} & \\ \hline & \frac{1}{m_1} & 0 \\ 0_{2\times 2} & 0 & \frac{1}{m_2} \end{array}\right). \tag{8.45}$$

Next, we evaluate if each of the diagonal matrices are positive definite. The lower right matrix is clearly positive definite as it is a diagonal matrix with strictly positive entries. For the upper left, a calculation is required

$$D^2V\big|_0 = \begin{pmatrix} V_1'' + V_2'' & -V_2'' \\ -V_2'' & V_2'' + V_3'' \end{pmatrix}\Bigg|_{x=0\in\mathbb{R}^4}. \tag{8.46}$$

Due to the fact that $V_1'' + V_2''|_0 > 0$ and $\det(D^2V|_0) > 0$ we have that $D^2V|_0$ is positive definite by Sylvester's theorem. Hence $D^2H$ is positive definite, i.e. cup shaped. Now we have shown that $H(x)$ is a Lyapunov function, and therefore guarantees that the origin is nonlinearly stable irrespective of the actual form of the potentials.

From this we wonder what else can be said about the global phase space geometry, which leads us to the Morse lemma.

*Lemma* 8.4 (Morse). *Let $x_0$ be a nondegenerate critical point for $f : \mathbb{R}^n \to \mathbb{R}$, i.e*

$$Df(x_0) = 0, \quad \det\left(D^2 f(x_0)\right) \neq 0. \tag{8.47}$$

*Let $k$ be the number of negative eigenvalues of $D^2 f(x_0)$. Then there exists a local change of coordinates (diffeomorphism) $x \mapsto y$ near $x_0$ such that in the new coordinates*

$$f(y) = f(x_0) - y_1^2 - \ldots - y_k^2 + y_{k+1}^2 + \ldots + y_n^2. \tag{8.48}$$

By the Morse lemma, there exists a change of coordinates $x \mapsto y$ such that $H(y) = y_1^2 + y_2^2 + y_3^2 + y_4^2$ that is suitable near the origin. Therefore the energy surface $E_{h_0} = \{y \in \mathbb{R}^4 : \ H(y) = h_0\}$ is equal to $\mathcal{S}^3 \subset \mathbb{R}^4$, where $\mathcal{S}^3$ is the 3-dimensional hyper-sphere. This means in the original $x$-coordinates, all low-energy level sets of $H(x)$ are diffeomorphic to 3-dimensional spheres.

It is possible to draw on other results from differential geometry, namely the foliation theorem [Milnor, 1965]. This tells us that level surfaces of a scalar function can only change their topology (i.e. become no longer diffeomorphic to each other) through level surfaces that contain a critical point of the function ($Df$ must vanish somewhere on the level set).

We can see this in the example of the pendulum equation, whose Hamiltonian we previously derived. The stable fixed point is a nondegenerate critical point of the Hamiltonian (as all other points around it have more energy) and by the Morse lemma, after a change of coordinates, level surfaces can be parametrized as a sum of quadratic terms. The loops representing periodic oscillation are also the level sets for the respective energy level and are all diffeomorphic to each other (they are all diffeomorphic to $\mathcal{S}^1$) until the level set which intersects the unstable fixed points. This transition is depicted in Fig. 8.6.

Regarding the two degree of freedom oscillator example previously handled, all surfaces are diffeomorphic to $\mathcal{S}^3$ as $x = 0$ is the only critical point of $H$.

## 8.2 Volumes under Hamiltonian flow maps

We would like to understand what happens with a volume of the phase space under the flow map of a Hamiltonian system. When we talk about the transformation of a volume $V(t_0)$ under the flow map, we mean to study the set $V(t)$ which results from applying the flow map to $V(t_0)$, i.e. $V(t) = F_{t_0}^t(V(t_0))$. The evolution of such a volume in shown in Fig. 8.7. However, before continuing to study this, we must prove Liouville's theorem.

**Theorem 8.5** (Liouville)**.** *For a dynamical system*

$$\dot{x} = f(x,t); \quad x \in \mathbb{R}^n; \quad f \in \mathcal{C}^1, \tag{8.49}$$

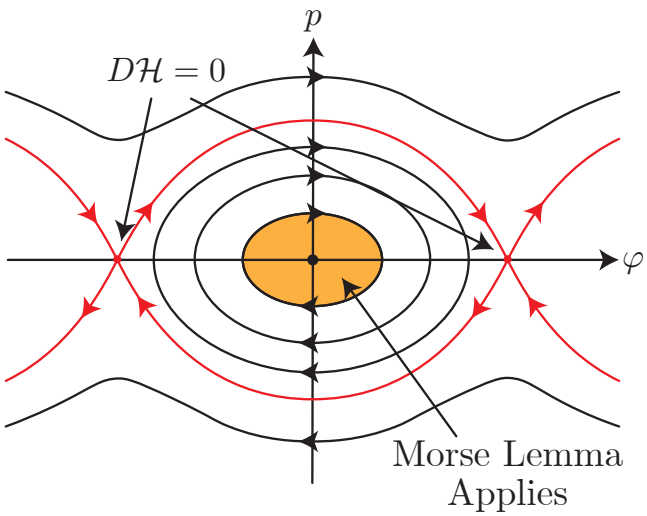


Figure 8.6: The effect of the foliation theorem. Around the stable fixed point where the Morse lemma applies is shaded orange, and the level sets containing critical points of $H$.

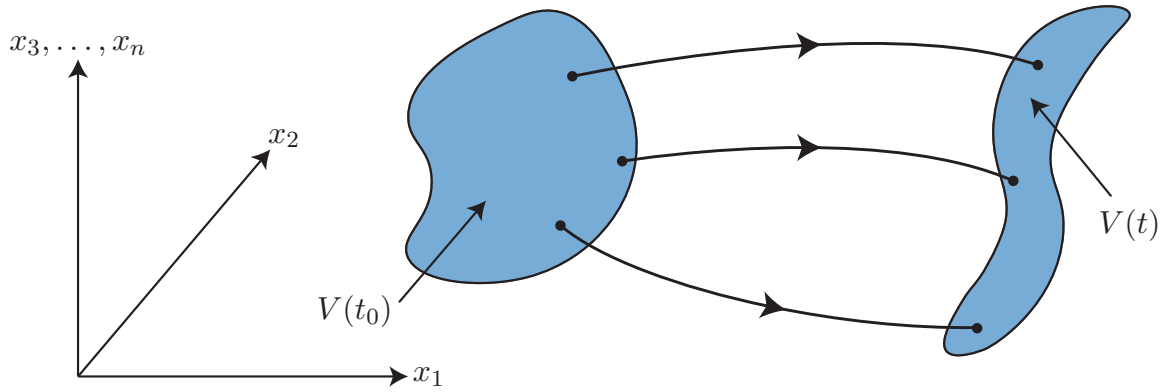


Figure 8.7: A transformed volume of the phase space under the flow map.

*we have that*

$$\boxed{\frac{d}{dt} vol(V(t)) = \int_{V(t)} \nabla \cdot f(x,t) dV.} \tag{8.50}$$

*Proof.* Calculate the equation of variations for this system with $x(t) = x(t; t_0, x_0)$

$$\frac{d}{dt}\frac{\partial x}{\partial x_0} = Df(x(t),t)\frac{\partial x}{\partial x_0} \implies \frac{d}{dt}\left[\frac{\partial x}{\partial x_0}\right] = \underbrace{Df(x(t),t)}_{A(t)}\frac{\partial x}{\partial x_0}. \tag{8.51}$$

This measures how sensitive the resulting $x(t)$ is on the initial condition $x_0$. We can generalize this to

$$\dot{X} = A(t)X; \quad A(t) = Df(x(t),t); \quad X, A \in \mathbb{R}^{n\times n}. \tag{8.52}$$

The only consistent initial condition for this is $\frac{\partial x}{\partial x_0}\Big|_{t=t_0} = X(t_0) = I_n \in \mathbb{R}^{n\times n}$. Recalling Abel's theorem, which applies to <u>any</u> linear system of ODE's we have

$$\det(\underbrace{X(t)}_{DF^t_{t_0}}) = \det(\underbrace{X(t_0)}_{I_n}) \cdot \exp\left(\int_{t_0}^{t} \mathrm{Tr}[A(s)]ds\right). \tag{8.53}$$

Note that in present case $\frac{\partial x}{\partial x_0} = DF^t_{t_0}$. Finally, for a general dynamical system we have

$$\det\left(DF^t_{t_0}\right) = \exp\left(\int_{t_0}^{t} \nabla \cdot f(x(s),s)ds\right). \tag{8.54}$$

Using this we calculate

$$\frac{d}{dt}\mathrm{vol}\left[V(t)\right] = \frac{d}{dt}\int_{V(t)} dV = \frac{d}{dt}\int_{V(t_0)} \det\left(DF^t_{t_0}(x_0)\right) dV_0 \tag{8.55a}$$

$$= \int_{V(t_0)} \frac{d}{dt}\det\left(DF^t_{t_0}(x_0)\right) dV_0 = \int_{V(t_0)} \nabla \cdot f(x(t),t) \underbrace{\det\left(DF^t_{t_0}\right) dV_0}_{=dV} \tag{8.55b}$$

$$= \int_{V(t)} \nabla \cdot f(x(t),t) dV. \tag{8.55c}$$

Where on the last equation of the first line we used the change of coordinates $y = F^t_{t_0}(x_0)$, then on the last equation of the second line we used our previous result, and then finally switched back to the original coordinates for the final equality. □

**Proposition 8.6** (Conservation of phase space volume)**.** *For any Hamiltonian system, the flow map $F^t_{t_0} : x_0 \mapsto x(t;t_0,x_0)$ is volume preserving for any $t_0, t \in \mathbb{R}$.*

*Proof.* For Hamiltonian systems we have that

$$\nabla \cdot (\underbrace{JDH(x,t)}_{f(x,t)}) = \mathrm{Tr}\left[JD^2H(x,t)\right] = \mathrm{Tr}\left[\begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}\begin{pmatrix} D^2_{qq}H & D^2_{qp}H \\ D^2_{qp}H & D^2_{pp}H \end{pmatrix}\right] \tag{8.56a}$$

$$= \mathrm{Tr}\left[\begin{pmatrix} D^2_{qp}H & * \\ * & -D^2_{qp}H \end{pmatrix}\right] = 0. \tag{8.56b}$$

Therefore by Liouville's theorem we have that $\mathrm{vol}[V(t)]$ is constant which is equivalent to $\det\left(DF^t_{t_0}(x_0)\right) = 1$. This is called this *infitesimal (local) form.* □

## 8.3 Consequences of area preservation

We now constrain ourselves to two dimensional, time-periodic Hamiltonian systems

$$\dot{x} = JDH(x,t); \quad x \in \mathbb{R}^2; \quad H(x,t) = H(x,t+T); \quad T > 0. \tag{8.57}$$

The Poincaré map is

$$P_{t_0} : x_0 \mapsto x(t_0 + T; t_0, x_0) = F_{t_0}^{t_0+T}(x_0). \tag{8.58}$$

$$H(x,t) = H(x_0) + \varepsilon H_1(x,t); \quad x = \begin{pmatrix} \varphi \\ p_\varphi \end{pmatrix}; \quad \dot{x} = JDH(x,t), \tag{8.59}$$

where $H_1$ and $H$ are $T$-periodic in $t$.

It is possible to classify the different types of fixed points for $P_{t_0}$. Let $x_0$ be a fixed point of $P_{t_0}$, i.e. $P_{t_0}(x_0) = x_0$, then we have by area preservation that

$$\det\left(DP_{t_0}|_{x_0}\right) = \det\left(DF_{t_0}^{t_0+T}(x_0)\right) = 1. \tag{8.60}$$

Hence, the eigenvalues $\lambda_1, \lambda_2 \in \mathbb{C}$ of the linearized Poincaré map have the property

$$1 = \det\left(DP_{t_0}(x_0)\right) = \lambda_1\lambda_2. \tag{8.61}$$

Furthermore, as the matrix $DP_{t_0}(x_0)$ is real and has even dimensions, the eigenvalues must occur in complex conjugate pairs. This implies that there are only a few restricted eigenvalue configurations which are possible.

(i) **Robust (structurally stable) configurations** Saddles with $\lambda_1 \neq \lambda_2$ and $\lambda_1, \lambda_2 \in \mathbb{R}$ and centers with $\lambda_1 \neq \lambda_2$ with $\lambda_1 = e^{i\alpha} = \overline{\lambda}_2$ are the only robust eigenvalue configurations which are possible.

(ii) **Non-robust configurations** The only other possibilities are $\lambda_1 = \lambda_2 = 1$ or $\lambda_1 = \lambda_2 = -1$.

The different eigenvalue configurations are shown in Fig. 8.8. As we can see the only robust fixed points of 2-dimensional Poincaré maps in Hamiltonian systems are saddles or centers (under linearization). The same conclusion holds, as a consequence, for any planar, autonomous Hamiltonian system (flow) $\dot{x} = JDH(x)$.

Next we examine what happens when we perturb homoclinic orbits. We recall the setting for homoclinic chaos in the Hamiltonian case with $H_1$ being $T$-periodic in $t$

$$\dot{x} = JD[H(x) + \varepsilon H_1(x,t)]. \tag{8.62}$$

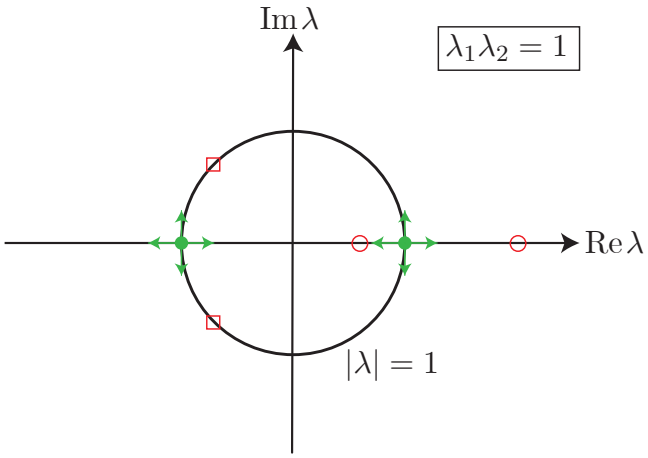


Figure 8.8: All possible eigenvalue configurations of the linearized Poincaré map.

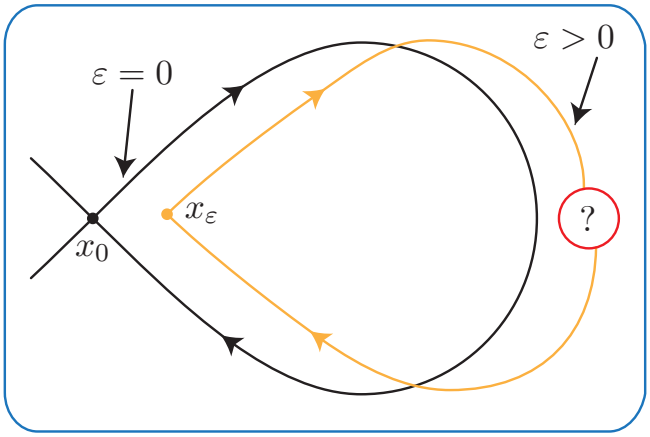


Figure 8.9: The unperturbed $\varepsilon = 0$ system with a homoclinic orbit in black and the perturbed $\varepsilon > 0$ system in yellow, the Melnikov method analyzes if the unstable and stable manifold intersect.

We assumed that the unperturbed system had a homoclinic obrbit and then asked if the unstable and stable manifolds of the perturbed system intersected using the Melnikov method, as sketched in Fig. 8.9.

**Proposition 8.7.** *The unstable manifold $W^U(x_\varepsilon)$ and stable manifold $W^S(x_\varepsilon)$ must intersect for $\varepsilon > 0$.*

*Proof.* Assume contrarilty that for $\varepsilon > 0$ the manifolds do not intersect. This leaves us with the image shown in Fig. 8.10.

As in Fig. 8.10, denote by $A$ a poin in $W^S(x_\varepsilon)$ and by $B$ a poin in $W^U(x_\varepsilon)$. Then call the

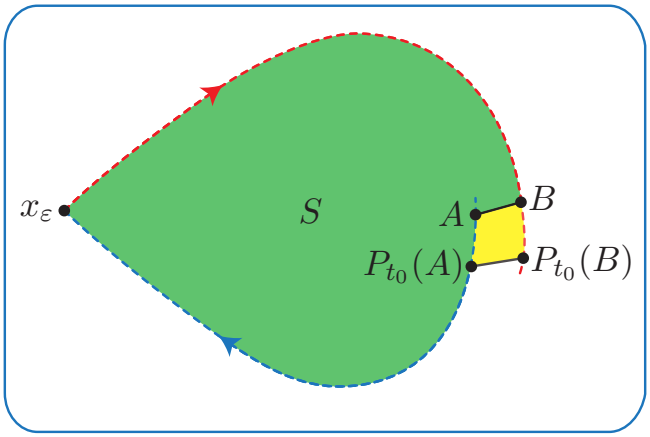


Figure 8.10: The image which we are left with if the two manifolds do not intersect.

region $S$ the area enclosed by the curves of each manifold and the line $AB$. The region $S$ under the Poincaré map now has a larger area than $S$, as can be seen by calculating:

$$\text{vol}[P_{t_0}(S) - S] = \text{vol}[\overline{ABP_{t_0}(A)P_{t_0}(B)}] \neq 0. \tag{8.63}$$

This violates the area preservation property of the Poincaré map, and therefore leads to a contradiction. □

There are four possible types of intersection geometries: transverse (chaotic), topologically transverse (chaotic), tangency, and equality (no chaos). These are all illustrated in Fig. 8.11. The tangency intersection lacks area preservation, and is also not possible.

An equality intersection can occur when the Hamiltonian is conserved even after perturbation. This implies that all trajectories continue to remain confined to the level sets of $H$.

As tangent intersections are not possible within the area preserving context, if $M(t_0) \neq 0$ then we must have homoclinic chaos. Thus in periodically forced Hamiltonian systems, chaos is generic! However, when we only have autonomous perturbations of the form $H(x) = H_0(x) + \varepsilon H_1(x)$, the equal case, in which the homoclinic orbit persists is generic.

**Proposition 8.8** (Recurrence)**.** *For an autonomous Hamiltonian system*

$$\dot{x} = JDH(x); \quad x \in \mathbb{R}^{2n}, \tag{8.64}$$

*suppose there exists a subset $S_0$ of $\mathbb{R}^{2n}$ which is compact, invariant, and has nonzero volume. Then for any open set $U \subset S_0$ and any time $T > 0$, there exists an $x_0 \in U$ and $t > T$ such that $x(t; t_0, x_0) \in U$.*

This property is illustrated in Fig. 8.12.

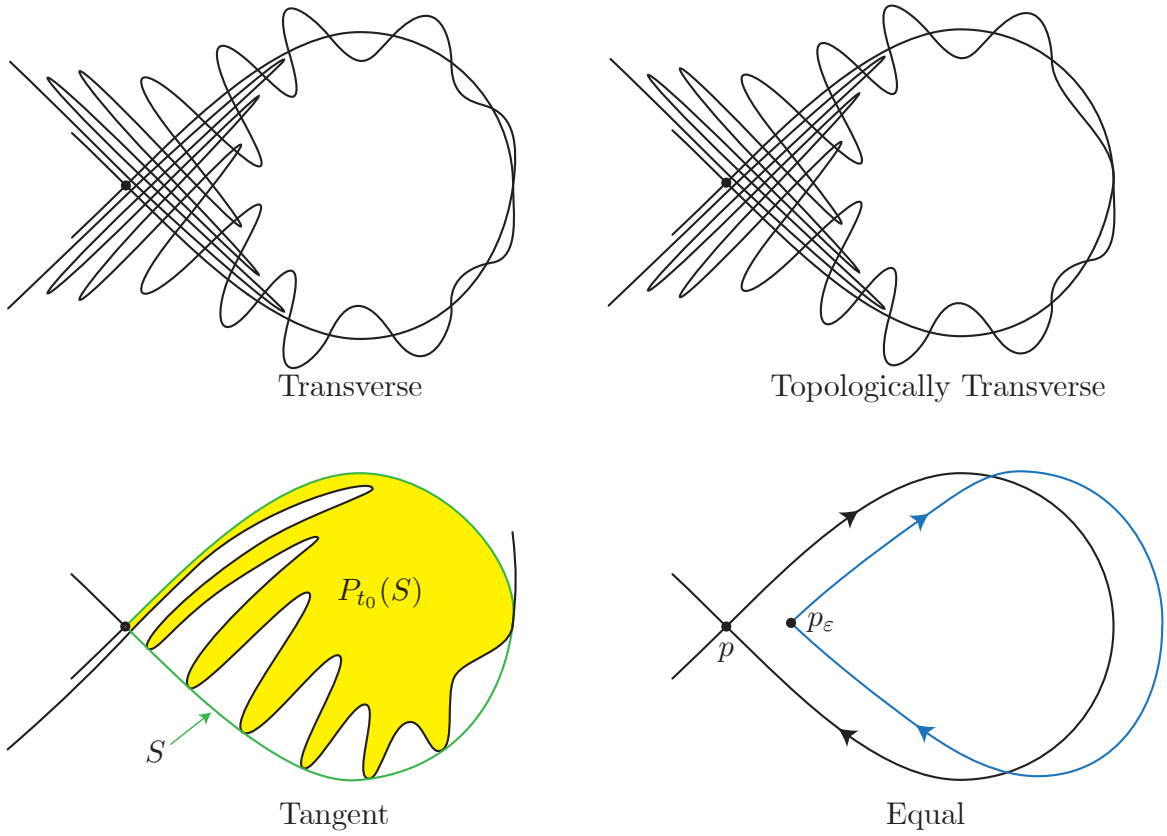


Figure 8.11: The four types of intersection geometries. Top row: transverse and topologically transverse, bottom row: tangent and equal.

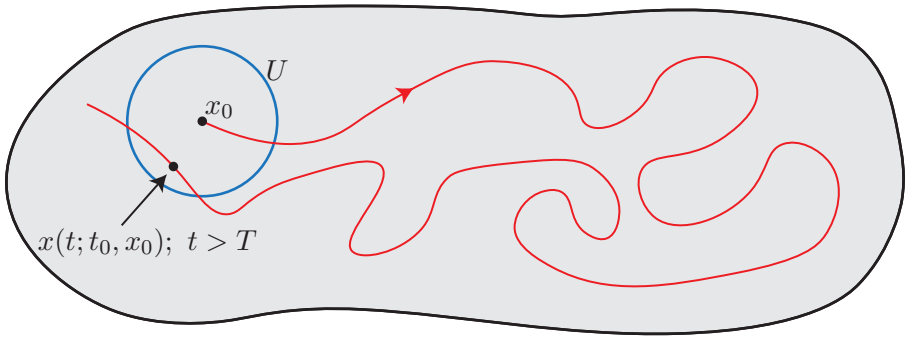


Figure 8.12: A depiction of the recurrence property of autonoumous, area-preserving, Hamiltonian systems.

*Proof.* The recurrence propery follows from the more general Poincaré Recurrence Theorem, which can be applied here as the flow map $F_{t_0}^t : x_0 \mapsto x(t;t_0,x_0)$ is volume preserving and

invertible and $S_0$ is compact, invariant, and has nonzero volume. □

**Proposition 8.9** (Poincaré Recurrence theorem)**.** *For a compact domain $B$ with nonzero volume equipped with an invertible, volume preserving map $P : B \to B$, for any open subset $U \subset B$ there exists a $N \in \mathbb{N}^+$ such that*

$$\boxed{P^N(U) \cap U \neq \emptyset.} \tag{8.65}$$

*Proof.* Consider the sequence $U, P(U), P^2(U), \ldots$, we claim that at least two of these sets intersect. Towards a contradiction, assume that each set is disjoint, therefore we have

$$\text{vol}[B] \geq \text{vol}\left[\bigcup_{i=0}^{\infty} P^i(U)\right] = \sum_{i=0}^{\infty} \text{vol}\left[P^i(U)\right] = \sum_{i=1}^{\infty} \text{vol}[U] = \infty. \tag{8.66}$$

The first inequality stems from the invariance of $B$, the second equality from the disjointness assumption, and the third from the volume preservation. This is clearly a contradiction as $B$ is compact and must have finite volume. Therefore there exist $m, n$ such that the intersection $P^m(U) \cap P^n(U)$ is non-empty. Call this intersection $A$ and assume without loss of generality that $n > m$.

For an $x \in A$ we know that $x \in P^m(U)$, thus $P^{-m}(x) \in U$, similarly we know $x \in P^n(u)$ so $P^{-m}(x) \in P^{n-m}(U)$. Putting these two facts together tells us that $P^{-m}(x)$ is an element of $U$ and $P^{n-m}(U)$, hence the intersection $U \cap P^{n-m}(U)$ cannot be empty. By calling $N = n - m$, we have shown the proposition. □

*Remark* 8.10. There is a stronger version of Poincaré's Recurrence theorem, which states that almost all initial conditions return arbitrarily close to themselves. Here almost all is with respect to the assumed volume.

*Example* 8.7 (Recurrence of the pendulum). For the pendulum, the area enclosed by an energy level set can be used as $S_0$. For large enough energy levels, this set may not be compact. This can be resolved by transforming the phase space $\mathbb{R}^2 \to \mathcal{S}^1 \times \mathbb{R}$ by taking advantage of the periodicity. This is shown in Fig. 8.13 where the importance of compactness and invariability are demonstrated. The figure also shows that one cannot simply apply the recurrence theorem to a Hamiltonian system restricted to an energy level, because this restricted system does conserve volume any more.

*Example* 8.8 (Recurrence in non-smooth conservative systems). For a linear spring with a mass attached, which can bounce off an elastic wall as the spring is expanded, we can also observe recurrence. Denoting the deviation of the mass by $q$ and the momentum by $m\dot{q}$ we still have recurrence under the map $P : x_0 \mapsto x(T; 0, x_0)$, which is volume preserving. The dynamical system and phase space are illustrated in Fig. 8.14.

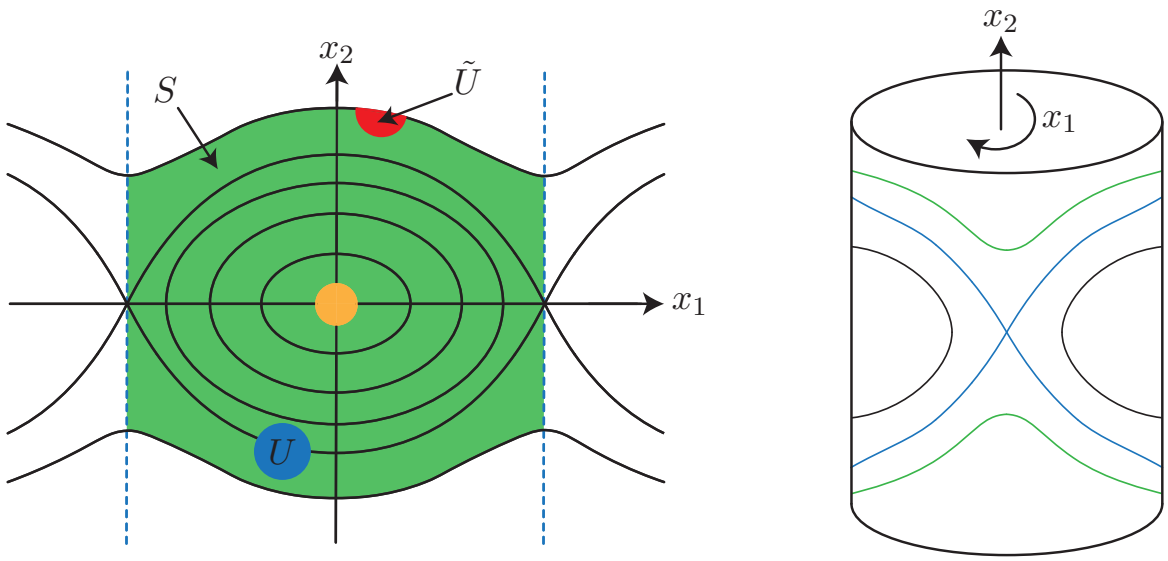


Figure 8.13: The recurrence of the pendulum is depicited along with the phase space transformation. The recurrence in the standard phase space does not hold for all points in the set $U$ (blue), because initial conditions below the separatrix never return. This is due to the fact that in the phase space $S_0$ is not compact. If instead we identify the two dotted blue lines to transform the phase space based on periodicity, we find that recurrence does hold. Let us now restrict the dynamics to the separatrix, resulting in a one-dimensional dynamical system. The set in the intersection of the separatrix and the set U. This is a relatively open and compact set which is also invariant for the restricted dynamics, yet these points never return. However, in this case, the dynamics is not volume preserving any more, which means we cannot apply the recurrence theorem.

*Remark* 8.11. These recurrent results extent to time-periodic Hamiltonian systems as well by applying them to the period-$T$ map. Furthermore, they can be extended to quasi-periodic Hamiltonians.

*Example* 8.9 (Ideal gas in a partitioned container). An ideal gas in a partitioned three dimensional container, which we denote by $B_0 \subset \mathbb{R}^3$. One partition is full of $n \gg 1$ gas molecules, and the other partition is a vacuum. After removing the partition, statistical reasoning leads to the conclusion that the molecules will spread evenly. The layout of the partition is shown in Fig. 8.15. The position of the of the particles is described by the variable $q = (x_1, y_1, z_1, \ldots, x_n, y_n, z_n)^T$ with $q \in B_0^n \subset \mathbb{R}^{3n}$, their respective velocities are described by $\dot{q} \in \mathbb{R}^{3n}$. The Lagrangian is made up of the kinetic energy $T(\dot{q}) = \frac{1}{2}\sum_{i=1}^{n} m_i \|v_i\|^2$ with $v_i = (\dot{x}_i, \dot{y}_i, \dot{z}_i)^T$ and the potential of the elastic interactions $V(q)$. With the generalized momentum given by $p = \frac{\partial L}{\partial \dot{q}}$, in this case the linear momentum, the Hamiltonian can be calculated

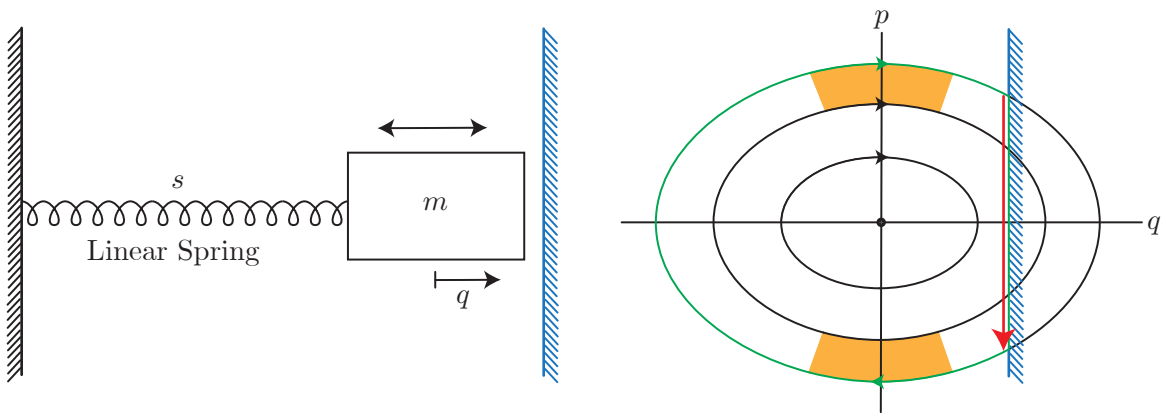


Figure 8.14: Left: A sketch of the dynamical system, with the elastic wall in blue. Right: The phase space of the dynamical system, with the red arrow designating that when the mass hits the wall (positive $x_2$) it immediately moves to the negative lower half (at the same absolute height) due to the elasticity. Furthermore, the volume preserving nature is shown by the two equal volume orange areas.

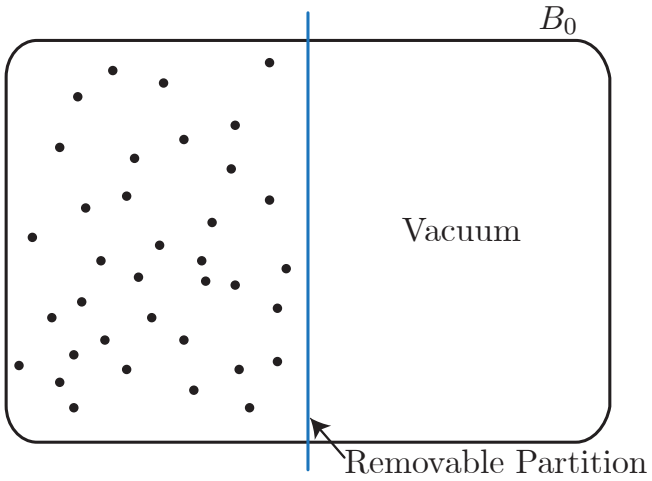


Figure 8.15: The gas contained in the container $B_0$ with the removable partition (blue) separating the gas and vacuum.

as $H = H(q,p) = H(x)$ for $x = (q,p)^T \in \mathbb{R}^{6n}$. At each collision there is a discontinuity of the velocity, however the flow map $F_{t_0}^t$ is still invertible and volume preserving.

The initial configuration of the gas is given by $x_0 \in \mathbb{R}^{6n}$. The unique solution $x(t)$ describes the complete evolution of all of the molecules (their positions and momenta). This is just a single trajectory in this (large) phase space $B_0^n \times \mathbb{R}^{3n} = \{q \in \mathbb{R}^{3n} : \ q_i \in [0,K]\} \times \mathbb{R}^{3n}$. An

example of such a path is depicted in Fig. 8.16. Observe that due to the conservation of energy

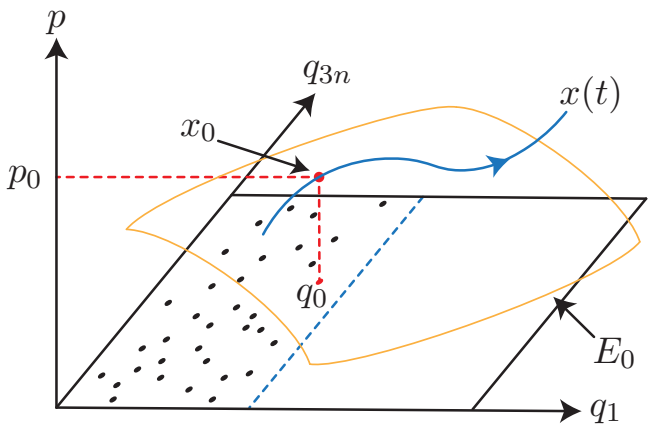


Figure 8.16: A single trajectory in the phase space for the partitioned container. Note that the trajectory is started from $x_0$ and remains within the energy level surface $E_0$.

we have that the following is conserved for all times

$$H_0 = H(x_0) = H(x(t)) = T(\dot{q}(t)) + V(q(t)). \tag{8.67}$$

Furthermore, the potential is always positive, i.e. $V(q) \geq 0$. This implies that $T(\dot{q}(t)) \leq H_0$, thereby bounding $\|\dot{q}_i(t)\|$ for all times $t$. Thus the momenta $\|p(t)\|$ are also bounded for all times as $p = m\dot{q}$. The positions $q_i(t)$ are bounded by construction (they must remain within the container), therefore the region $S_0 = \{x \in \mathbb{R}^{6n} : H(x) \leq H_0\}$ is compact and invariant. By recurrence, almost all initial gas configurations return arbitrarily close to themselves over time. However, the recurrence time would be comparable to the life of the universe.

*Remark* 8.12. We cannot apply the recurrence to the energy surface as the restricted map may not be volume preserving.

The next example is of some historical value as this was the first ever numerical experiment to test a theory.

*Example* 8.10 (Fermi-Ulam-Pasta experiment (1955)). The experiment was conducted for large systems of linear, undamped, coupled oscillators. Since the system is linear, the energy content of each of the normal modes cannot change in time, and hence there is no equipartition of energy. According to the conclusions from statistical mechanics, a weak nonlinearity should change this completely and distribute the energy equally among the modes.. Initially, the experiments confirmed this hypothesis. However, after forgetting to turn off their computer for the night, they observed repeated returns close to the original energy configuration in their

simulation, with the energy practically localized again in the same mode (ca. 3% error). More and more close returns were observed over longer time intervals.

The reason for this is that in Hamiltonian systems energy surfaces are compact and therefore recurrence applies for the phase space volume contained inside the energy surfaces.

## 8.4 Stability of fixed points in Hamiltonian systems

Recall the setup for autonomous Hamiltonian systems

$$\dot{x} = JDH(x), \quad x \in \mathbb{R}^{2n}. \tag{8.68}$$

The fixed points of such systems are critical points of the Hamiltonian, i.e. $DH(x_0) = 0$. Observe therefore that no asymptotically stable fixed points may exist in a Hamiltonian system. This is due to the volume preservation property. Assuming contrarily that the fixed point $x_0$ is asymptotically stable implies that there exists a ball $B_0$ within the open domain of attraction such that $F_{t_0}^t(B_0)$ is strictly contained within $B_0$ for $t$ large enough. This would require that $\text{vol}[F_{t_0}^t(B_0)] < \text{vol}[B_0]$ for such a $t$, clearly violating the volume preservation property. Thus, at best, $x_0$ can be Lyapunov stable, and the nonlinear terms are relevant.

We first explore what we can find from the linear stability. To this end let $y = x - x_0$ and calculate the linearized system at $x_0$

$$\dot{y} = JD^2H(x_0)y. \tag{8.69}$$

Note that the linearized system is also Hamiltonian

$$\tilde{H}(y) = \frac{1}{2}\langle y, D^2H(x_0)y \implies \dot{y} = JD\tilde{H}(y). \tag{8.70}$$

Thus the phase space volume is also preserved for the linear system. This can be shown as we have before with

$$\nabla \cdot JD\tilde{H}(y) = \nabla \cdot JD^2H(x_0)y = \text{Tr}[JD^2H(x_0)] = 0. \tag{8.71}$$

This also tells us that the sum of the eigenvalues of the linearized Hamiltonian system is zero.

Further, observe that if $\lambda$ is an eigenvalue of (8.69) (whether real or complex), then so is $-\lambda$. This leaves us with a few possible eigenvalue configurations for Hamiltonian systems, as seen in Fig. 8.17.

Suppose $\lambda = 0$ is an eigenvalue, then it is necessarily at least a double eigenvalue. Furthermore, under a variation of the parameters, purely imaginary eigenvalues can only leave the imaginary axis through a collision with another eigenvalue. Recalling that a linearized

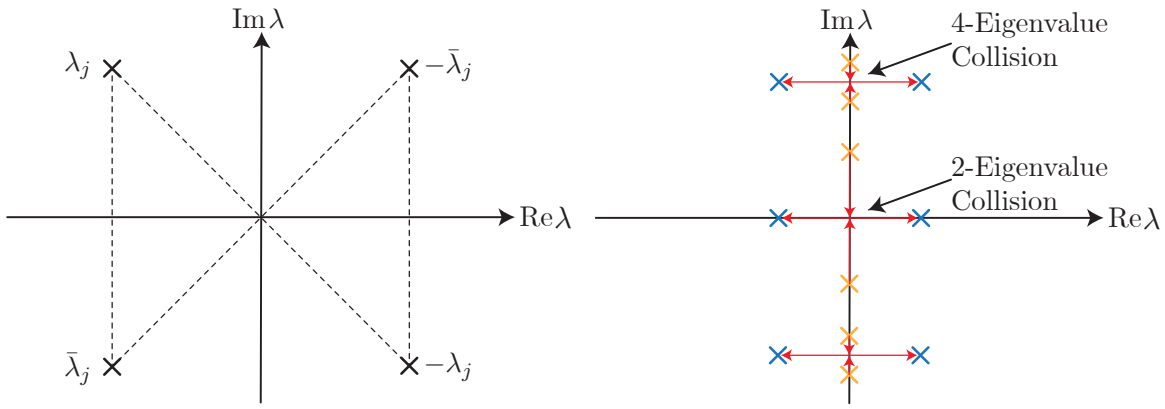


Figure 8.17: Left: An example eigenvalue configuration with the constraints found previously. Right: The possible losses of stability through eigenvalue collisions

Hamiltonian can only be neutrally stable leads to the conclusion that linear stability analysis, is never conclusive to establish Lyapunov stability in Hamiltonian systems.

The previous examination leads us to focusing on the nonlinear stability of fixed points of Hamiltonian. The setup is the same as before with an autonomous Hamiltonian system underlying the dynamics. Recall that for such systems we have that $\frac{dH(x(t))}{dt} = 0$, i.e. the Hamiltonian is conserved under the flow map, and as before a fixed point $x_0$ corresponds directly to a critical point of the Hamiltonian, i.e. $DH(x_0) = 0$.

These properties of $H$ make it a candidate for a Lyapunov function. Observe that if $D^2H(x_0)$ is positive definite, then it is cup shaped and choosing $H$ as a Lyapunov function implies that $x_0$ is Lyapunov stable. Meanwhile, if $D^2H(x_0)$ is negative definite, then choosing $-H$ as a Lyapunov function also implies that $x_0$ is stable. Hence, if $D^2H(x_0)$ is definite then $x_0$ is nonlinearly Lyapunov stable.

*Example* 8.11 (Multiple degree of freedom nonlinear mechanical system). Consider three pendulums, with masses $m_1$, $m_2$, and $m_3$ attached in order by nonlinear springs. The mass $m_3$ is attached to a mass $m_4$ which is constrained to move on a loop of radius $R$. The position of each mass is measured by its displacement from the unstretched spring position by the variable $q$, thus the springs are in their unstretched positions at the position $q = 0$. The system is depicted in Fig. 8.18.

Concatenating the position and momentum into the variable $x$ we have $x = (q, p)^T \in \mathbb{R}^8$. It is clear that $DH(0) = 0$ and due to the total mechanical energy having a strict minimum at $x = 0$ we have that $D^2H(0)$ is positive definite. Thus $x = 0$ is Lyapunov stable.

We may wonder if all linearly stable Hamiltonian fixed points are in fact always nonlinearly

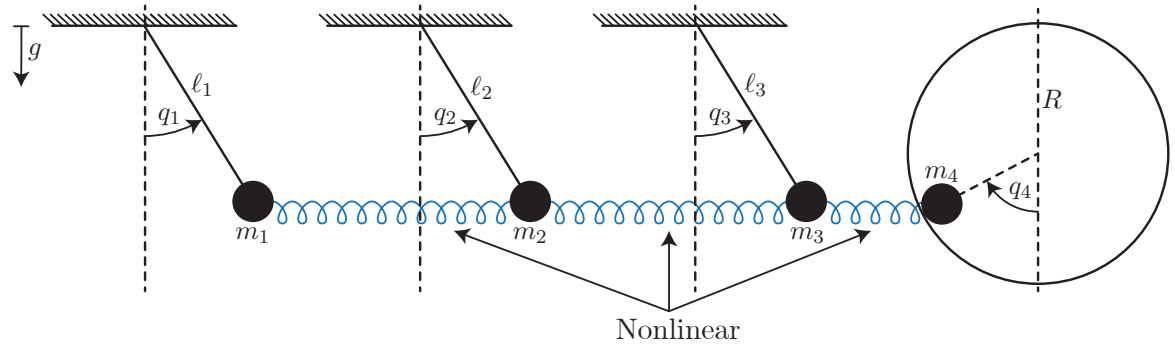


Figure 8.18: A depiction of the multiple degree of freedom nonlinear dynamical system consisting of three pendulums and a loop.

stable, i.e. if $D^2H(x_0)$ is always definite. This is not the case, as we can see by taking

$$H = \frac{1}{2}\left(q_1^2 + p_1^2\right) - \frac{1}{2}\left(q_2^2 + p_2^2\right); \quad x = \begin{pmatrix} q_1 \\ p_1 \\ q_2 \\ p_2 \end{pmatrix} \in \mathbb{R}^4. \tag{8.72}$$

This leads to the dynamics

$$\begin{cases} \dot{q}_1 = p_1 \\ \dot{q}_2 = -p_2 \\ \dot{p}_1 = -q_1 \\ \dot{p}_2 = q_2 \end{cases} \implies \dot{x} = \left(\begin{array}{cc|cc} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{array}\right) x. \tag{8.73}$$

The eigenvalues of this system are $\pm i$, thus linear stability analysis and use of $H$ as a Lyapunov function (see Fig. 8.19) are both inconclusive.

*Example* 8.12 (Gyroscopic systems). A classic example is that of a gyroscopic system such as the tippe-top toy (a spinning top like a dreidel). When spun there is a relative equilibrium fixed point for the Hamiltonian obtained by factoring out the cyclic coordinate, which is achieved via a Routh-reduction of cyclic coordinates. Fast spinning stabilizes this Hamiltonian equilibrium, but the Hamiltonian has no minimum there (nonlinear analysis is needed to prove this). Additional dissapation will then always destabilize these equilibria in a process called *dissapation-induced instability* which is shown in Fig. 8.20.

With these examples, it is clear that further study of nonlinear stability of Hamiltonian fixed points is necessary for the case that $D^2H(x_0)$ is not definite. The setting we will work in

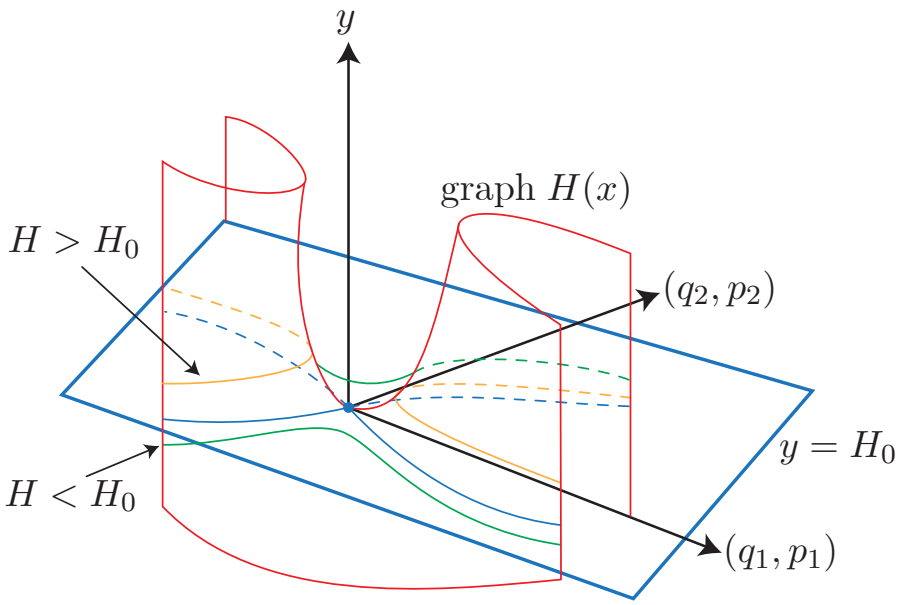


Figure 8.19: The graph of the Hamiltonian $H$ showing the lack of definiteness which leads to an inconclusive Lyapunov stability analysis.

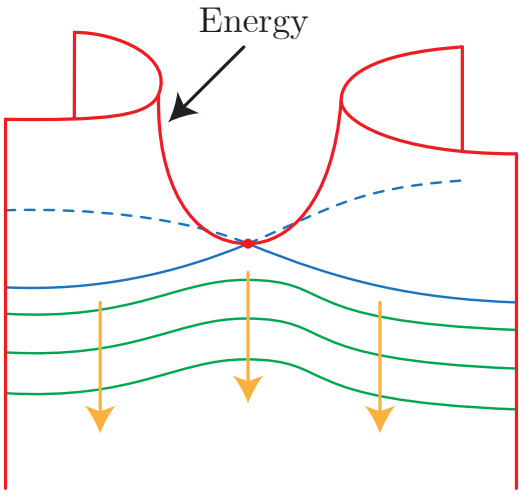


Figure 8.20: The energy is shown in red, the dissipation of energy takes the motion towards lower energy levels as shown by the yellow arrows.

to study this is as follows

$$\dot{x} = JDH(x); \quad x \in \mathbb{R}^{2n}; \quad DH(0) = 0. \tag{8.74}$$

The final equality is always achievable by a coordinate shift. Once again we begin by linearizing

to find $\dot{y} = Ay$. Assume for simplicity that $A$ is semisimple, i.e. it has $2n$ linearly independent eigenvectors (even if some eigenvalues are repeated). Through a change of coordinates $A$ can be block-diagonalized

$$A = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_n \end{pmatrix}; \quad A_j = \begin{pmatrix} 0 & \omega_j \\ -\omega_j & 0 \end{pmatrix}. \tag{8.75}$$

Each block $A_j$ corresponds to the eigenvalues $\pm i\omega_j$. There is no real part to the eigenvalue, otherwise $x = 0$ would be unstable, however $\omega_j < 0$ is possible. This change of coordinates allows us to frame the geometry of the linearization as the product of $n$ different two dimensional dynamical systems, each with clockwise rotating cyclical orbits, i.e. $\prod_{i=1}^{n} \mathcal{S}^1 = \mathbb{T}^n$ the $n$-dimensional torus. This product is shown in Fig. 8.21.

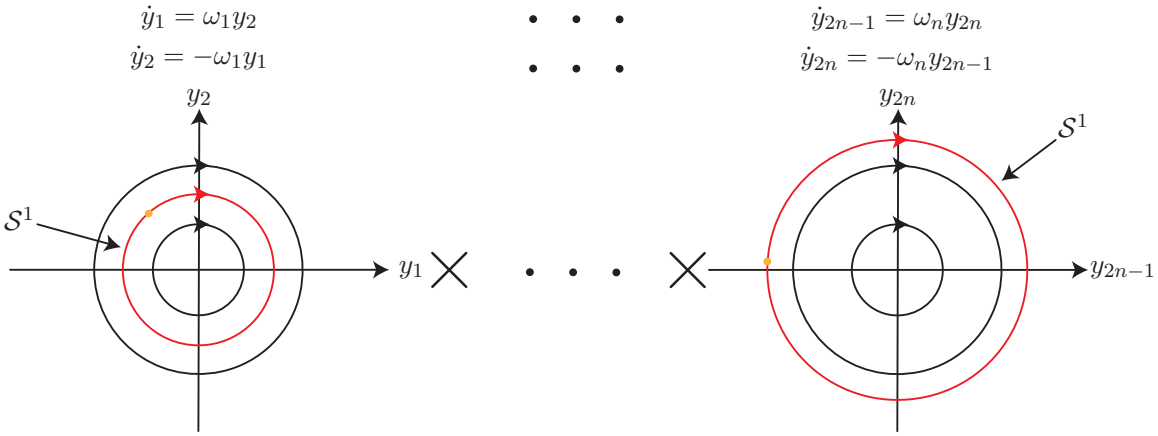


Figure 8.21: The product of the $n$ dynamical systems of the linearization forming the $n$-dimensional torus.

Defining the vector of angular velocities $\omega = (\omega_1, \ldots, \omega_n)$ we get three cases for the structure of the orbits

(i) **Resonance** If there exists a $k \in \mathbb{Z}_+^n$ (i.e. $k_i \neq 0$ for all $i$) with $\langle k, \omega \rangle = 0$, trajectories close up. Therefore all orbits are closed and periodic on $\mathbb{T}^n$.

(ii) **Non-resonance** If for each $k \in \mathbb{Z}^n$ with $\|k\| \neq 0$ we have that $\langle k, \omega \rangle \neq 0$, trajectories never close up. Therefore all orbits are quasi-periodic and yield a dense winding.

(iii) **Partial resonance** If there exists a $k \in \mathbb{Z}^n$ with $\|k\| \neq 0$ and $\langle k, \omega \rangle = 0$ we get a combination of the previous two cases. Trajectories fill lower dimensional tori densely and are not closed.

## 8.5 Stability of Hamiltonian fixed points: normal forms

The setting of the previous section remains the same as we move to introduce the idea of the normal form of $H(x)$ near $x = 0$. This requires two additional assumptions

(i) For each $k \in \mathbb{Z}^n$ with $\|k\|_1 = \sum_{i=1}^n |k_i| \leq 4$ we have $\langle k, \omega \rangle \neq 0$, i.e. there are no resonances below order 5.

The first assumption implies that a nonlinear, near-identity canonical change of coordinates puts $H(x)$ in the form

$$H(\varphi, r) = \sum_{i=1}^{n} \omega_i r_i + \frac{1}{2} \sum_{k,l=1}^{n} \omega_{k,l} r_k r_l + \mathcal{O}\left(|r|^4\right); \quad r \in \mathbb{R}^n_{\geq 0}; \quad \varphi \in \mathbb{T}^n. \tag{8.76}$$

The $\varphi$-dependence starts only in the higher order terms, i.e. $\mathcal{O}\left(|r|^n\right)$ in this case. By a near-identity change of coordinates, we mean that $(r, \varphi)$ is close to the polar coordinates introduced for the linearized Hamiltonian.

**Definition 8.1.** A *canonical* (also *symplectic*) coordinate change is one which preserves the Hamiltonian structure. That is, in the new coordinates, Hamilton's equations still hold

$$\begin{cases} \dot{\varphi} = \dfrac{\partial H}{\partial r} \\ \dot{r} = -\dfrac{\partial H}{\partial \varphi}. \end{cases} \tag{8.77}$$

These reductions bring us to the next theorem. In order to state this, some terminology must be rigorously defined for this section.

**Definition 8.2.** *Persists* in this cases corresponds to surviving the transition back from the linearization to the nonlinear system through a smooth deformation by a small amount.

**Definition 8.3.** *Most* (in this case invariant tori) means

$$\lim_{\varepsilon \to 0} \frac{\mu(\text{persisting tori in } \varepsilon\text{-ball})}{\mu(\mathcal{B}(\varepsilon))} = 1, \tag{8.78}$$

where $\mu$ is the measure that is preserved, i.e. the phase space volume.

(ii) The *nondegeneracy condition* holds, i.e. $\det[\omega_{kl}] \neq 0$.

**Theorem 8.13** (Kolmogorov-Arnold-Moser (KAM) theorem)**.** *We have three items which hold:*

*(i) Under the assumptions (i) and (ii), most invariant tori of the linearized Hamiltonian persist in the nonlinear system close to $x = 0$.*

*(ii) Surviving toi are quasi-periodic, implying that trajectories are dense on these and obey ODEs of the form*

$$\begin{cases} \dot{r} = \mathcal{O}\left(|r|^4\right) \\ \dot{\varphi} = \tilde{\omega}(r, \varphi) = \dfrac{\partial H}{\partial r} = \omega + [\omega_{kl}]r + \mathcal{O}\left(|r|^3\right). \end{cases} \tag{8.79}$$

*(iii) Under the additional nondegeneracy condition known as the* isoenergentic nondegeneracy *condition*

$$\det \left[\begin{array}{c|c} [\omega_{kl}] & \omega^T \\ \hline \omega & 0 \end{array}\right] \neq 0, \tag{8.80}$$

*most tori persist within each energy surface close enough to the origin. This is shown in Fig. 8.22.*

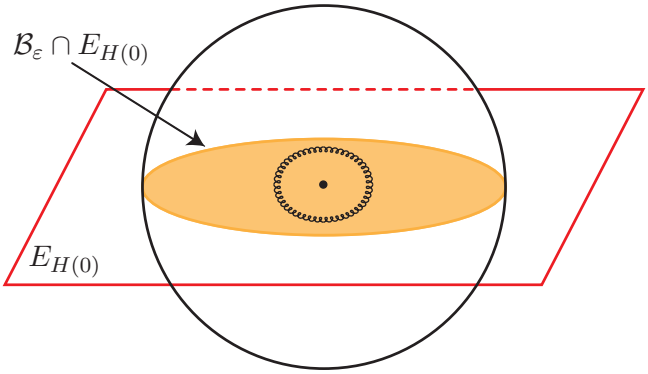


Figure 8.22: Most tori persist with respect to the measure defined inside the energy surface $E_{H(0)}$ (red). The $\varepsilon$-ball is given in black, and its intersection with $E_{H(0)}$, which is of dimension $2n - 1$.

This has different implications for the dynamics near $x = 0$ depending on the degrees of freedom based on the dimension of $E_{H(x)}$ $(= 2n - 1)$. For $x \neq 0$ close to the origin the dimension of the surviving tori is dim $\mathbb{T}^n = n$, thus the codimension of $\mathbb{T}^n$ within $E_{H(x)}$ is $(2n - 1) - n = n - 1$.

(i) **1 degree of freedom** (e.g. pendulum) The codimenension of $\mathbb{T}^1$ is 0, therefore the torus is the energy surface itself and $x = 0$ is Lyapunov stable.

(ii) **2 degrees of freedom** (e.g. two coupled pendulums) The codimension of $\mathbb{T}^2$ is 1, so the two-dimensional torus locally divides the three-dimensional energy surface into an interior and exterior. The interior is a compact invariant set that trajectories cannot escape from within the given energy surface. Hence $x = 0$ is Lyapunov stable, even if $D^2H(0)$ is indefinite.

(iii) $\geq$ **3 degrees of freedom** (e.g. three coupled pendulums) The codimension of $\mathbb{T}^n$ is at least 2, thus the tori no longer form barriers and trajectories can sneak out. Generically, trajectories do sneak out along a chaotic invariant set formed near the remnants of destroyed (resonant) tori such as in *Arnold diffusion*.

In Fig. 8.23 we show a nested family of surviving KAM-tori and the internal structure of the gaps between them. The gaps are due to resonances and are filled with lower dimensional hyperbolic and elliptic tori. A heteroclinic tangle forms as the intersection of the stable and

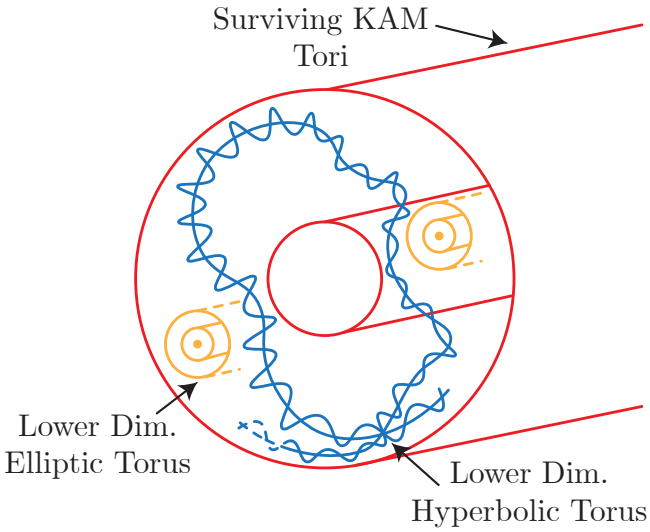


Figure 8.23: The geometry of Arnold diffusion. The red tori are two surviving KAM-tori of a large dimensional dynamical system. In the gaps between them we find hyperbolic and elliptic tori, which are lower dimensional.

unstable manifolds of the hyperbolic (saddle-type tori). This tangle works as a transition chain and can produce $\mathcal{O}(1)$ displacements in $r$ for trajectories started near the surviving torus, even when $\varepsilon \to 0$.

Therefore we have a mechanism for instability when $D^2H(0)$ is not definite. When it is definite, the transition chains are still there but energy surfaces are small spheres near the origin. This implies that complicated dynamics does not lead to instability.

**Theorem 8.14** (Nekhoroshev's theorem)**.** *Under further nondegeneracy conditions, the speed of diffusion is less than* $C \exp\left(-\frac{1}{\varepsilon^d}\right)$*, where* $d$ *depends on* $H(x)$*. This means that the speed is exponentially small.*

# Chapter 9

# Normally hyperbolic invariant manifolds

The objective of this chapter is to extend the theory of low-dimensional stable and unstable manifolds to a general, higher-dimensional setting. The focus will lie on invariant manifolds rather than individual trajectories as invariant manifolds are more observable, more influential, and more robust. We begin with an example.

*Example* 9.1 (Dynamics near a saddle-type fixed point). The difference between the behavior of individual trajectories and of the invariant manifolds is depicted in Fig. 9.1. The idea being that under a diffeomorphic transformation of the invariant manifolds of the linear system into those of the nonlinear system, a trajectory starting from a point $x_0$ may be drastically different even if the dynamics remain topologically conjugate to each other. I.e. individual trajectories are sensitive with respect to initial conditions and changes in parameters, meanwhile robust invariant manifolds persist (perturb smoothly).

Thus our examination now turns to determining which invariant manifolds are persistent. In order to understand this, first an intermezzo in differentiable manifolds is necessary.

## 9.1 A crash course on differentiable manifolds

**Definition 9.1.** A set $M \subset \mathbb{R}^n$ is a *$k$-dimensional differentiable manifold*, if it is locally diffeomorphic to $\mathbb{R}^k$, i.e. for every $x \in M$ there exists an open (in $M$) neighborhood $V \subset M$, such that $V$ is diffeomorphic to an open set $U \subset \mathbb{R}^k$. The diffeomorphism is called the *local parameterization* as is denoted $\Phi : U \to V$. Its smooth inverse is the *local coordinate system* $\Phi^{-1} : V \to U$. These are illustrated in Fig. 9.2.

The intuitive meaning of a $k$-dimensional manifold is a set which locally looks like $\mathbb{R}^k$.

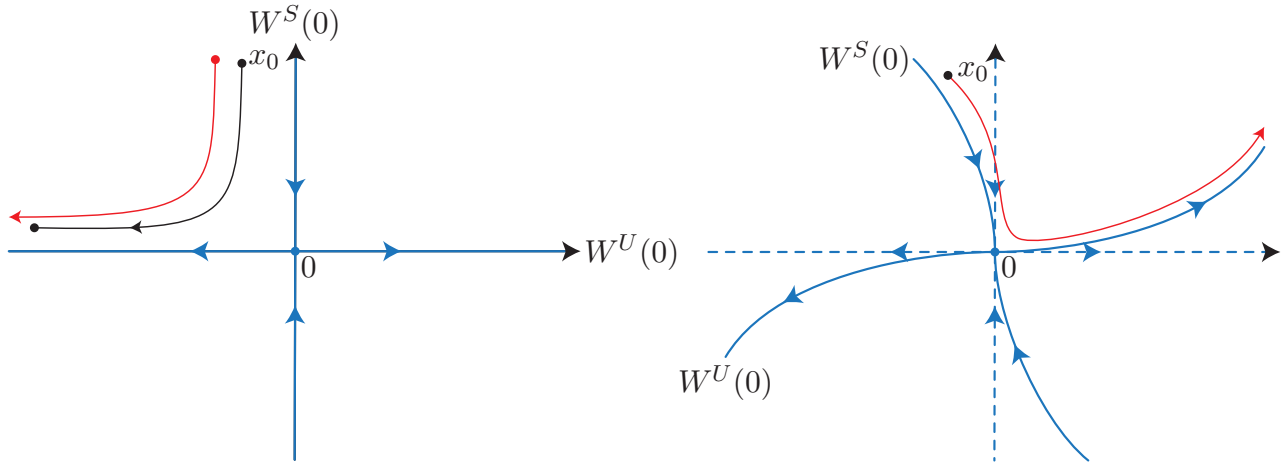


Figure 9.1: The dynamics near a saddle-type fixed point demonstrating the sensitive dependence of individual trajectories on initial conditions and changes in parameters, and the robust nature of invariant manifolds.

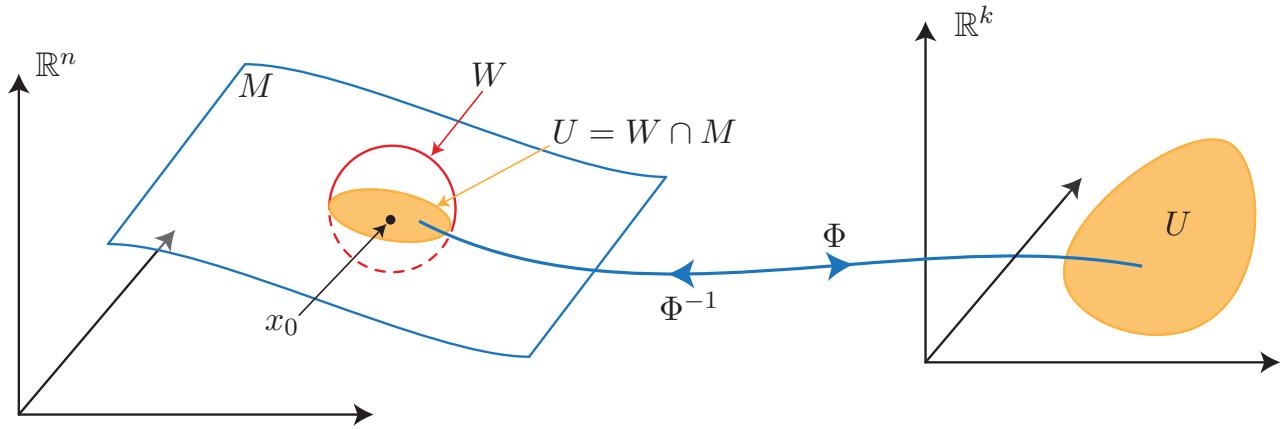


Figure 9.2: The differentiable manifold $M$ is given in blue in the left. The red sphere $W$ is an open ball in $\mathbb{R}^n$, intersecting this with $M$ yields and open (in $M$) neighborhood $V = W \cap M$ of $x$. The smooth diffeomorphism $\Phi$ transforms $U$ into the open set $U$ of $\mathbb{R}^k$.

**Definition 9.2.** For a set $M \subset \mathbb{R}^n$ and an open set $U \subset \mathbb{R}^n$, the set $M \cap U$ is called *relatively open* in $M$.

*Example* 9.2 (The circle is a 1-dimensional manifold). The circle $\mathcal{S}^1$ is a 1-dimensional differentiable manifold. To see define the four open subsets $V_i$ of $\mathcal{S}^1$ as open semicircles which are each shifted by $\frac{i\pi}{2}$ for $i = 1, 2, 3, 4$. Then $\mathcal{S}^1 = \bigcup_{i=1}^4 V_i$. These are drawn in Fig. 9.3. Any $x \in \mathcal{S}^1$ belongs to at least one of the $V_i$ due to the previous equality. Furthermore each $V_i$ is

diffeomorphic to $U = (-1, 1)$ which is open in $\mathbb{R}^1$ via the diffeomorphisms

$$\Phi_1 = \begin{pmatrix} t \\ \sqrt{1-t^2} \end{pmatrix}; \qquad \Phi_2 = \begin{pmatrix} t \\ -\sqrt{1-t^2} \end{pmatrix}; \tag{9.1a}$$

$$\Phi_3 = \begin{pmatrix} -\sqrt{1-t^2} \\ t \end{pmatrix}; \qquad \Phi_4 = \begin{pmatrix} \sqrt{1-t^2} \\ t \end{pmatrix}. \tag{9.1b}$$

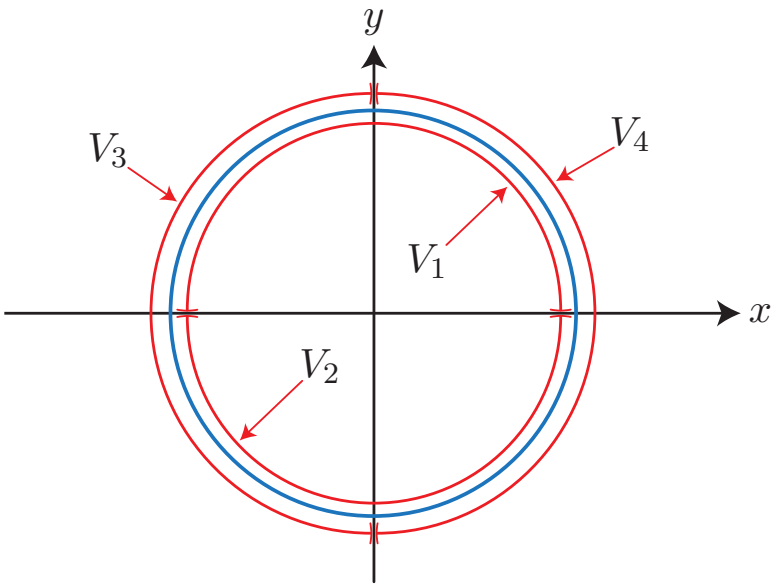


Figure 9.3: The subsets $V_i$ of $\mathcal{S}^1$ which are used to show that the circle is a 1-dimensional manifold.

There is no global parameterization of $\mathcal{S}^1$, but this is not needed to be a differentiable manifold, and these local parameterizations suffice.

*Example* 9.3 (Not all sets are manifolds). A few examples of sets which are not manifolds.

(i) The figure-eight is not a manifold, as for any point except the central intersection, there exists a local neighborhood diffeomorphic to $\mathbb{R}^1$, however at this central intersection, no open neighborhood exists which is diffeomorphic to $\mathbb{R}^1$. This is depicted in Fig. 9.4.

(ii) A dense orbit on a 2-dimensional torus is not a manifold, as around any point infinitely many adjacent trajectories exist, but do not form a plane (there is a distance between two instances of the trajectory). Hence any neighborhood around any point cannot be hyperbolic to $\mathbb{R}^2$ or $\mathbb{R}^1$. Such a neighborhood is illustrated in Fig. 9.5.

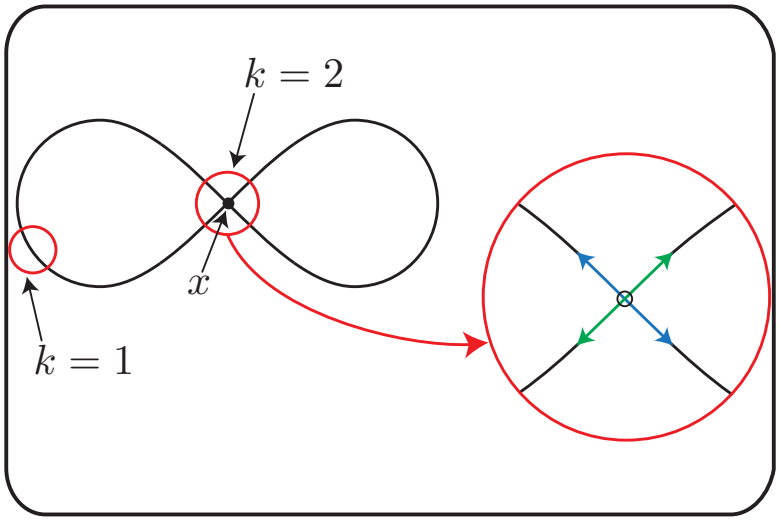


Figure 9.4: The figure-eight with the critical intersection which prevents it from being a manifold.

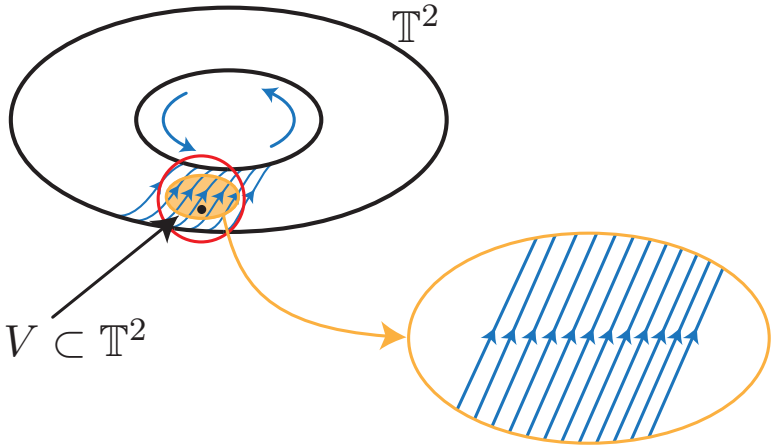


Figure 9.5: A depiction of a local neighborhood of a point in a dense orbit of the torus.

(iii) The union of a semi-infinite spiral and semi-open (left closed, right open) interval do not form a differentiable manifold. For this set, at the intersections of the two sets, points with neighborhoods as in the figure eight appear.

We can construct manifolds very easily, for instance by taking any smooth function $f : X \to Y$, the graph of $f$ namely the set $\text{graph}(f) = \{(x, y) : \ x \in X, \ y = f(x)\}$ is a differentiable manifold.

**Definition 9.3.** For two manifolds $X$ and $Z$, each subsets of $\mathbb{R}^n$, with $X \subset Z$ we call $X$ a

*submanifold* of $Z$.

An example of a submanifold would be $X = \mathcal{S}^1 \subset Z = \mathcal{S}^2 \subset \mathbb{R}^3$. We say $\mathcal{S}^1$ is a 1-dimensional submanifold of the 2-dimension manifold $Z$.

**Definition 9.4.** A manifold $M \subset \mathbb{R}^n$ is a *$k$-dimensional manifold with boundary*, if it is locally diffeomorphic to a relatively open set in the non-negative half-space $H^k = \{x \in \mathbb{R}^k : x_1 \geq 0\} \subset \mathbb{R}^k$. The *boundary of $M$*, denoted $\partial M$ is always mapped into $\partial H^k$ under any parameterization, and is a $k-1$-dimensional manifold. Such a manifold and its parameterization are illustrated in Fig. 9.6.

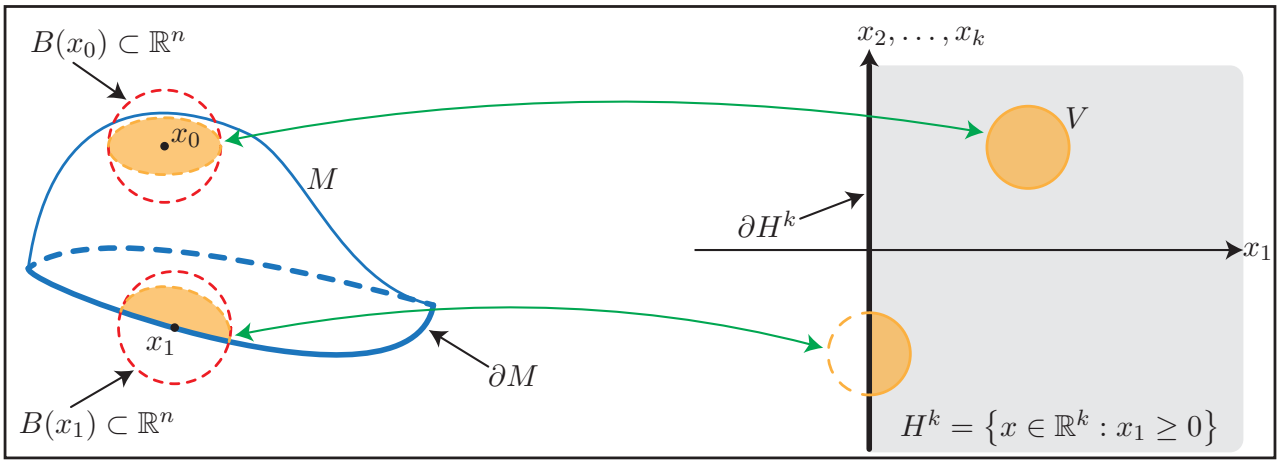


Figure 9.6: An illustration of the manifold with boundary $M$ and a parameterization mapping its boundary to $\partial H^k$.

Closed intervals, closed cyclinders, and closed hollow spheres are all manifolds with boundaries, however it is not always the case that simply taking the closure of a manifold yields a manifold with boundary. For instance the interior of a rectangle is a manifold (using the identity as a parameterization), however as the corners are not smooth the closure is not a manifold.

**Definition 9.5.** Given a $k$-dimensional manifold $M \subset \mathbb{R}^n$ with the local parameterization at $x$ given by $\Phi :: U \subset \mathbb{R}^k \to V \subset M \subset \mathbb{R}^n$ (assume for simplicity that $\Phi(0) = x \in V$), the *tangent space of the manifold* is then defined as

$$\boxed{T_x M = \text{Range}[D\Phi(0)].} \tag{9.2}$$

The intuition as to why this is called the tangent space is show in Fig. 9.7.

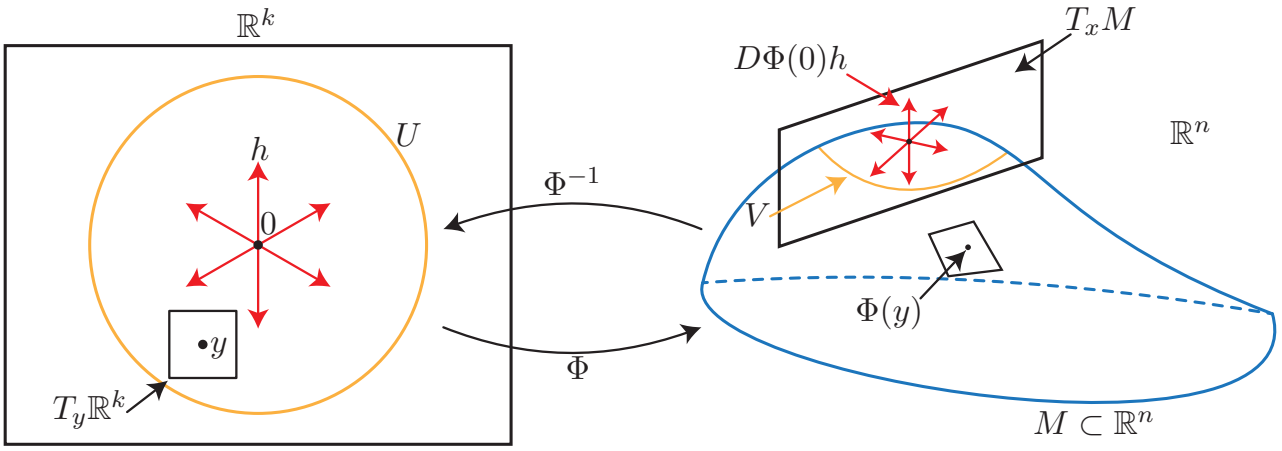


Figure 9.7: The tangent space of a manifold $M$. The map $D\Phi(0)$ maps the left image to the plane indicated in the right image.

Note here how the differential is defined

$$D\Phi(x)h = \lim_{s\to 0} \frac{\Phi(x+sh) - \Phi(x)}{s}. \tag{9.3}$$

Therefore the map $D\Phi(y)$ is a map from the tangent space of $\mathbb{R}^k$ at $y$ to the tangent space of $M$ at $\Phi(y)$, i.e. $D\Phi(y) : T_y\mathbb{R}^k \to T_{\Phi(y)}M$.

*Remark* 9.1. Note a few facts about tangent spaces:

(i) Although it appears that $T_xM$ depends on the parameterization $\Phi$, it actually does not;

(ii) The dimension of the tangent space, which is well defined as it is a linear subspace of $\mathbb{R}^n$, is equal to the dimension of $M$, namely $k$;

(iii) A map $f : M \to N$ between manifolds induces a linear mapping between the appropriate tangent spaces of the manifolds. Denoting by $\Phi$ and $\Psi$ the respective local parameterizations, the induced map between parameter spaces is $F = \Psi^{-1} \circ f \circ \Phi$. This is a mapping between linear spaces and $DF$ is well-defined. In turn, this implies that $Df$ is well-defined and by the chain rule is equal to $D\Psi \circ DF \circ D\Phi^{-1}$, therefore it is possible to differentiate mappings between manifolds. Then we have the sought-after mapping between tangent spaces

$$Df : T_xM \to T_{f(x)}N. \tag{9.4}$$

These maps are illustrated in Fig. 9.8.

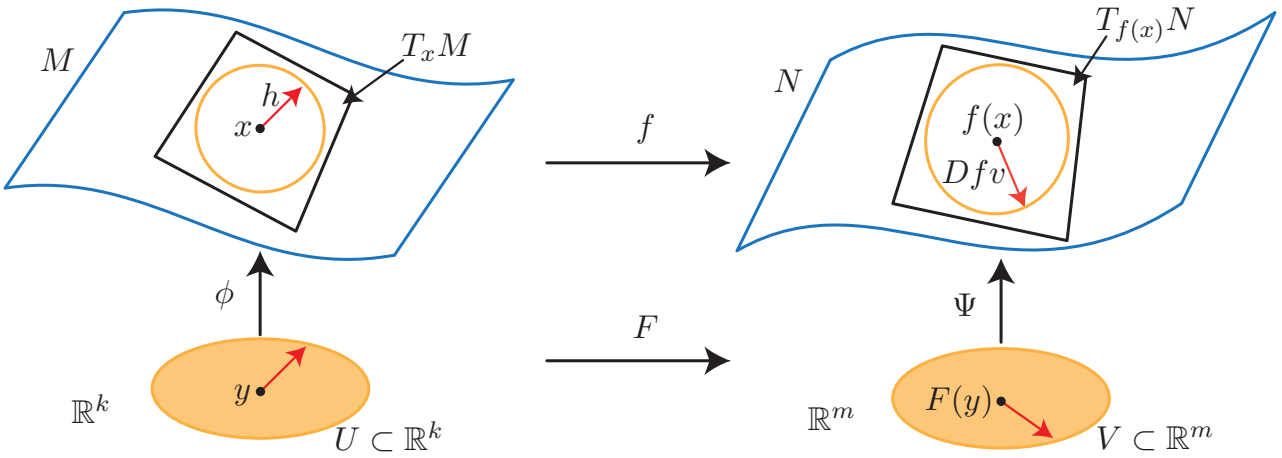


Figure 9.8: The commutative diagram along with the respective spaces illustrated for the induced map between tangent spaces.

**Definition 9.6.** The set of all tangent spaces, labelled by their base points is called the *tangent bundle*

$$\boxed{TM = \{(x, v) : \ x \in M, \ v \in T_xM\} = \bigcup_{x \in M} (x, T_xM).} \tag{9.5}$$

The tangent bundle is a $2k$-dimensional manifold.

By labelling tangent vectors with their respective base points, different tangent spaces are separate *fibers* of the tangent bundle. In general $TM \neq M \times \mathbb{R}^k$, i.e. the tangent bundle is not a global direct product. In fact, locally, each tangent bundle is trivial (the direct product), this just does not extend globally. However, for $M = \mathcal{S}^2$ the tangent bundle is diffeomorphic to $\mathcal{S}^2 \times \mathbb{R}^2$.

**Definition 9.7.** The *normal space*, unlike the tangent space, depends on the ambient space. For a $k$-dimensional manifold $M$ embedded in $\mathbb{R}^n$, the normal space at a point $x$, $N_xM$, is the orthogonal complement of the tangent space. For their direct sum it holds, that

$$\boxed{T_xM \oplus N_xM = T_x\mathbb{R}^n.} \tag{9.6}$$

Therefore, the dimensional of the normal space is $n - k$. The normal space is drawn in Fig. 9.9.

**Definition 9.8.** Analog to the tangent bundle, the *normal bundle* is the collection of normal spaces of the manifold $M$ labelled by their base points

$$\boxed{NM = \{(x, v) : \ x \in M, \ v \in N_xM\}.} \tag{9.7}$$

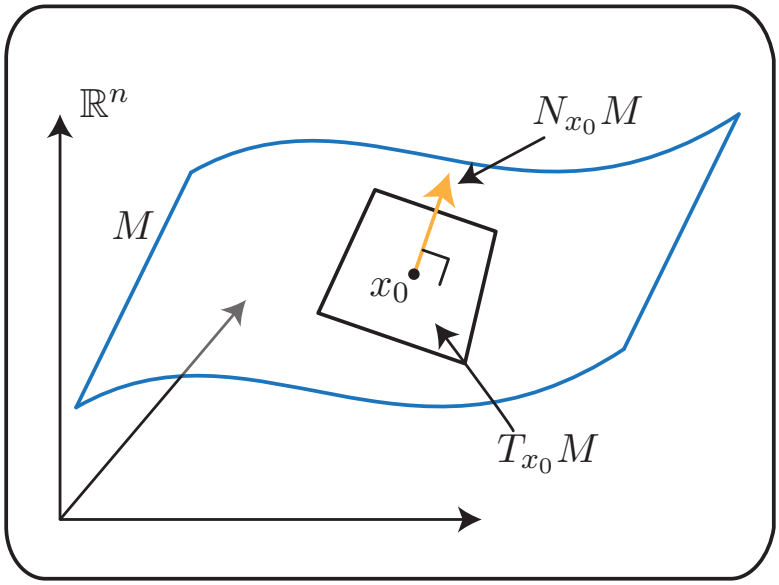


Figure 9.9: The normal space of a manifold $M$ at the point $x_0$ along with the tangent space at the same point.

By labelling normal vectors with their base points, intersections are avoided as shown in Fig. 9.10.

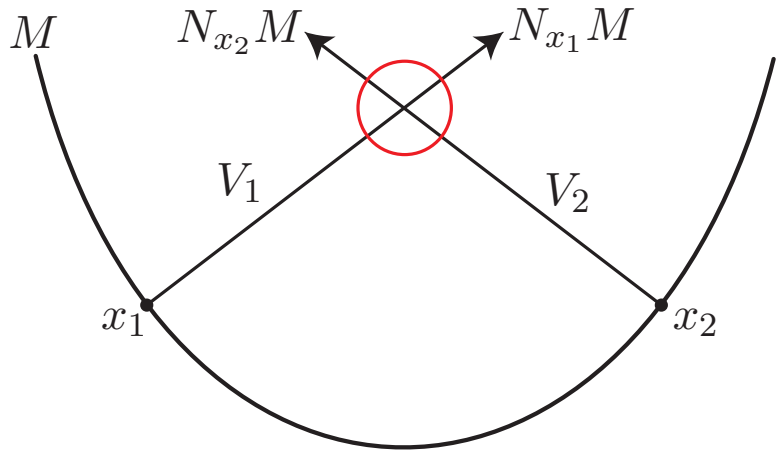


Figure 9.10: Intersecting normal vectors, these are avoided by labelling vectors with their base points.

**Definition 9.9.** A *subbundle* of the normal bundle is any $x$-dependent family of subspaces of $N_xM$, which is smooth in $x$.

An example of such a subbundle is given in Fig. 9.11, where the smooth family of subspaces is given by $S(x) \subset N_x M$.

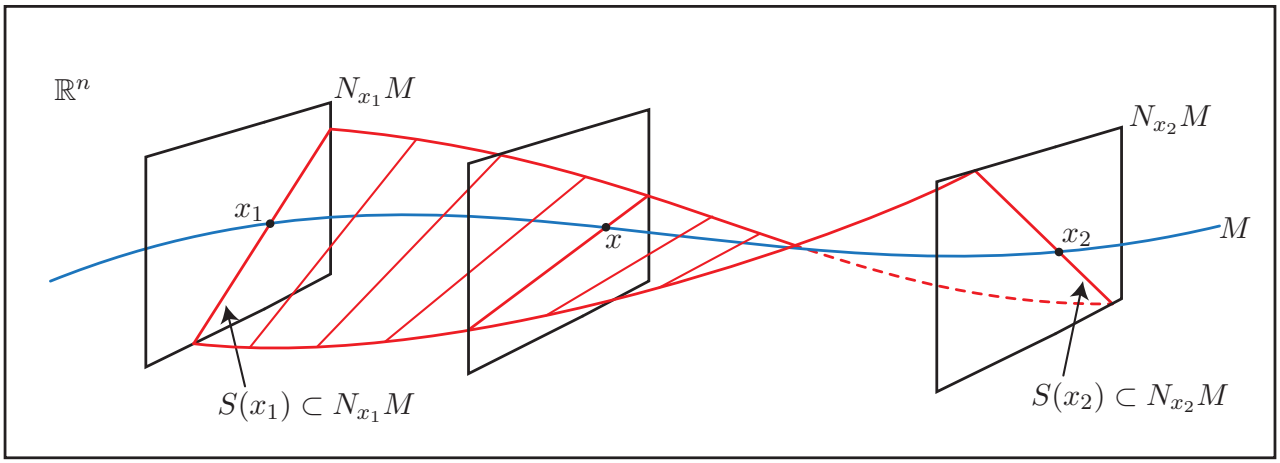


Figure 9.11: A subbundle of the manifold $M$, each base point $x \in M$ is a base point of $S(x)$, a fiber of the normal bundle.

## 9.2 Theory of normally hyperbolic invariant manifolds

The starting point for the following section will be of an autonomous dynamical system

$$\dot{x} = f(x); \quad x \in \mathbb{R}^n; \quad f \in \mathcal{C}^r \; r \geq 1. \tag{9.8}$$

Furthermore, let $M_0 \subset \mathbb{R}^n$ be a $k$-dimensional, compact, $\mathcal{C}^1$-manifold with boundary $\partial M_0$.

**Definition 9.10.** The manifold $M_0$ is an *overflowing invariant manifold* if $f(x)$ points strictly outwards of $\partial M_0$. A special case of this is when $\partial M_0 = \emptyset$.

There are three central questions for what happens to this manifold which we will address:

(i) Does it persists (survive smoothly)?

(ii) Is it stable (what happens to nearby trajectories)?

(iii) What is the interior structure of the stable/unstable manifolds of $M_0$?

*Example* 9.4. Consider the system

$$\begin{cases} \dot{x} = -x(1-x)(1+x) \\ \dot{y} = \alpha y \end{cases}; \quad 0 < \alpha \leq 1. \tag{9.9}$$

Now define $M_0$ to be the closed interval $[-1-\varepsilon, 1+\varepsilon] \times \{0\}$. This manifold is of dimension 1, has nonempty boundary $\{-1-\varepsilon, 1+\varepsilon\}$, is compact, and smooth. At each point in the boundary, $f$ is pointing away from $M_0$. Therefore, $M_0$ is an overflowing invariant manifold. We also have that $W^S_{\text{loc}}(M_0)$ is nonempty, meanwhile $W^U_{\text{loc}}(M_0)$ is empty. Trajectories of this dynamical system along with $M_0$ and the stable local manifold are depicted in Fig. 9.12.

Now we may ask if $M_0$ is robust under perturbations (if it persists). For example under the perturbation

$$\begin{cases} \dot{x} = -x(1-x)(1+x) - \varepsilon y \\ \dot{y} = -\alpha y + \varepsilon x \end{cases}; \quad 0 \leq \varepsilon \ll 1. \tag{9.10}$$

The fixed points $A_0$, $B_0$, and $C_0$ are hyperbolic, therefore they persist. In fact $B_0$ does not move at all. The other fixed points move to new perturbed counterparts, i.e. $A_0, C_0 \to A_\varepsilon, C_\varepsilon$. From the second equation, we see that the y coordinate of the perturbed fixed point becomes $y_\varepsilon = \varepsilon/\alpha x_\varepsilon$. That is, $C_\varepsilon$ moves upwards and $A_\varepsilon$ moves downwards.

At $B_0 = B_\varepsilon$ we can linearize to find

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} -1 & -\varepsilon \\ \varepsilon & -\alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \tag{9.11}$$

Depending on $\alpha$ we differentiate two cases.

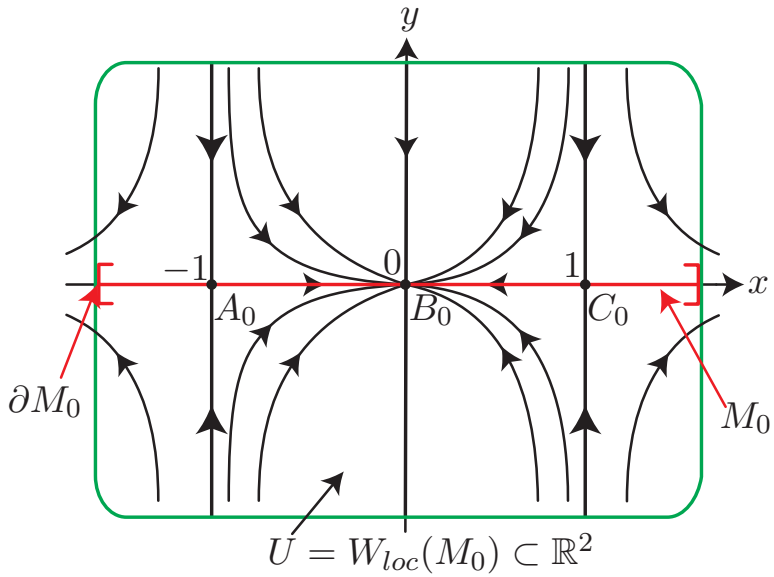


Figure 9.12: An illustration of the trajectories of the dynamical system, the overflowing invariant manifold $M_0$, its boundary, and the local stable manifold drawn in green. The fixed points $A_0$, $B_0$, and $C_0$ are also shown.

(i) $\boldsymbol{\alpha = 1}$ In this case the eigenvalues split in a vertical fashion, as shown in Fig. 9.13. The stable node turns into a stable spiral, which, when connected to the perturbed fixed points $A_\varepsilon$ and $C_\varepsilon$ cannot be diffeomorphic to $M_0$.

(ii) $\boldsymbol{\alpha < 1}$ Now change the perturbation to $\varepsilon^2$ such that both fixed points move upwards, the perturbed curve will have a cusp at 0, so we have again no diffeomorphism to $M_0$.

Thus in each case, no smooth transformation to $M_0$ exists and therefore there is no persistence. The perturbed geometries are shown in Fig. 9.14.

The reason for the destruction of the manifold here was the normal rate of attraction being less than or equal to the tangential rate of attraction at one point. This can be exemplified as follows.

*Example* 9.5. Consider the dynamical system

$$\begin{cases} \dot{x} = -x(1-x)(1+x) + \varepsilon y \\ \dot{y} = -\dfrac{3}{2}y - \varepsilon x \end{cases}. \tag{9.12}$$

This is the specification of the previous example to $\alpha = \frac{3}{2}$. If we linearize at the point $B_0$ for

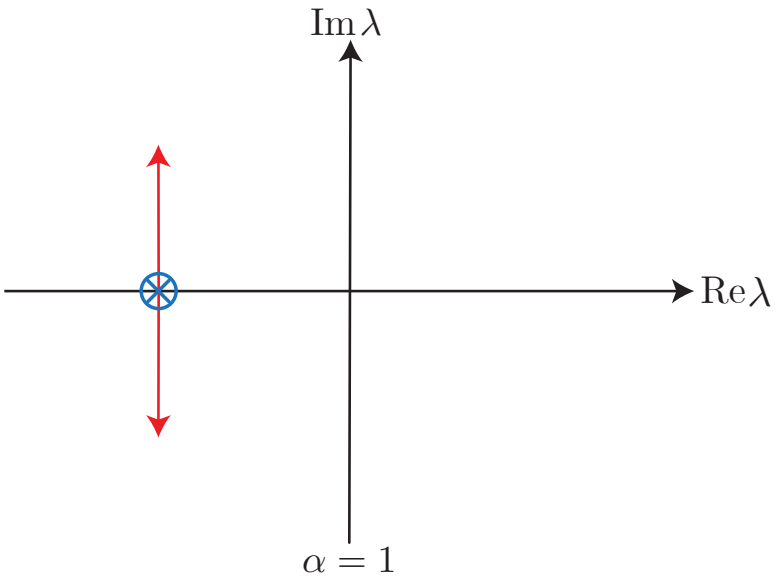


Figure 9.13: The eigenvalues splitting in a vertical fashion for $\alpha = 1$.

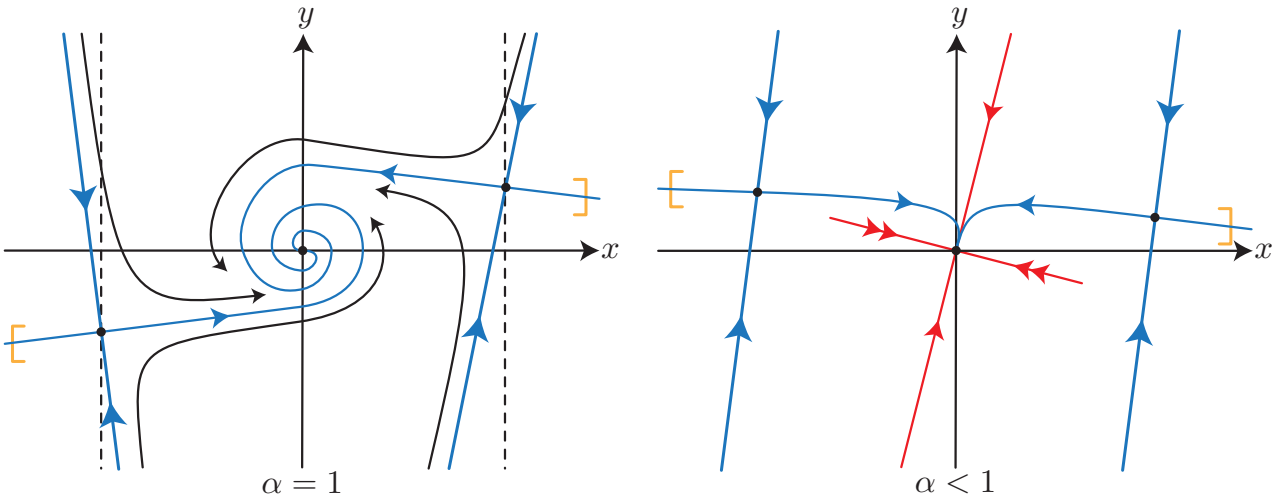


Figure 9.14: Left: The geometry for $\alpha = 1$. Right: The geometry for $\alpha < 1$.

$\varepsilon = 0$ we find

$$\begin{cases} \dot{x} = -x \\ \dot{y} = -\dfrac{3}{2}y \end{cases}. \tag{9.13}$$

The solution to the linearized system is

$$\begin{cases} x(t) = x_0 e^{-t} = x_0 \left(e^{-1}\right)^t \\ y(t) = y_0 e^{-\frac{3}{2}t} = y_0 \left(e^{-\frac{3}{2}}\right)^t \end{cases} \implies y = y_0 \left(\frac{x}{x_0}\right)^{\frac{3}{2}}. \tag{9.14}$$

The eigenvalue constellation and linearized phase portrait, along with the implied nonlinear phase portrait for this system are show in Fig. 9.15. From (9.14) we see that the phase curves of the system are only $\mathcal{C}^1$ because of the power of $\frac{3}{2}$. In the nonlinear system, this implies that the perturbed manifold $M_\varepsilon$ exists, but is also only $\mathcal{C}^1$-smooth. Thus, the ratio between normal and tangential contraction determines the smoothness of $M_\varepsilon$.

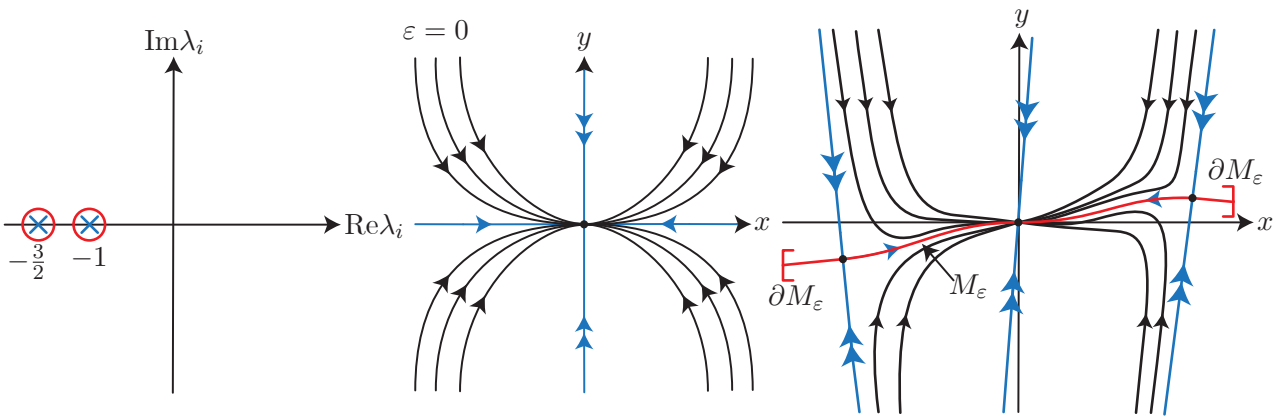


Figure 9.15: Left: The eigenvalue constellation for the linearization around $B_0$ for $\varepsilon = 0$. Middle: The linearized phase portrait around $B_0$. Right: The nonlinear phase portrait with the nearby overflowing invariant manifold $M_\varepsilon$.

One can check here that $\dot{y} = (r+\mu)y + \varepsilon x$ would give $\mathcal{C}^r$ differentiability for $M_\varepsilon$ for $0 \leq \mu \ll 1$. Formally, for $\mu = 0$ we would have $\mathcal{C}^\infty$ differentiability, but only for the linear part, and this would not be robust.

The general challenge is to quantify the growth/contraction rates in the normal and tangential directions, i.e. how do we determine the rates away from fixed points along trajectories.

Recall our setup with the compact, overflowing invariant $\mathcal{C}^r$ manifold $M_0 \subset \mathbb{R}^n$ of dimension $k$, for $r \geq 1$ and the dynamical system

$$\dot{x} = f(x); \quad x \in \mathbb{R}^n; \quad f \in \mathcal{C}^r. \tag{9.15}$$

Now consider the *inverse linearized flow map* on the tangent bundle of $\mathbb{R}^n$ for some $t \geq 0$

$$(F^{-t}, DF^{-t}) : \ T\mathbb{R}^n \to T\mathbb{R}^n; \ (p, v) \mapsto \left(F^{-t}(p), DF^{-t}(p)v\right). \tag{9.16}$$

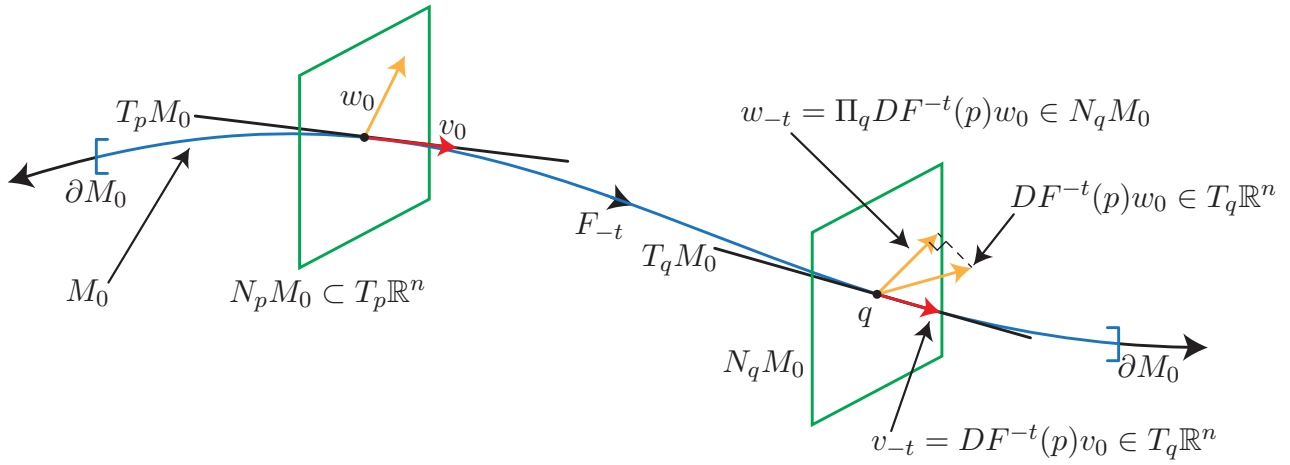


Figure 9.16: The inverse linearized flow map. We have $q = F^{-t}(p) \in M_0$ and the red arrow on the right is $v_{-t} = DF^{-t}(p)v_0 \in T_qM_0$ for $v_0 \in T_pM_0$. The map $\Pi_q$ is the orthogonal projection onto $N_qM_0$.

This map is depicted in Fig. 9.16.

To quantify the rate of normal attraction we define for $w_0 \in N_pM_0$, as in Fig. 9.16, $w_{-t} \in N_qM_0$ to be $\Pi_q DF^{-t}(p)w_0$. If we have repulsion in backward time, $\|w_{-t}\| \gg \|w_0\|$, with an exponential rate, i.e.

$$\frac{\|w_0\|}{\|w_{-t}\|} \sim e^{-\alpha t} = (e^{-\alpha})^t; \quad \alpha > 0; \quad t \to \infty, \tag{9.17}$$

then the best estimate for such an $e^{-\alpha}$ is

$$\boxed{\nu(p) = \inf\left\{a \in \mathbb{R}^+ : \lim_{t\to\infty} \frac{\|w_0\|}{\|w_{-t}\|}/a^t = 0, \ \forall w_0 \in N_pM_0\right\}.} \tag{9.18}$$

The inf is defined as the largest lower bound. The idea behind this, is that for large $a$, the quotient will go to zero, if we start to lower $a$ until this trend changes. We then pick the largest lower bound for this change.

If $e^{-\alpha} = a \sim \nu(p) < 1$ then $M_0$ is exponentially attracting in backward time. Note that the orbit from $p$ will generally not approach a fixed point. Furthermore, $\log(\nu(p))$ will generally not be the eigenvalue of a Jacobian (the linearized flow map at a fixed point).

Next, we would like to quantify the relationship between the normal and tangential rates. If we have

$$\frac{\|w_0\|}{\|w_{-t}\|}/\frac{\|v_0\|}{\|v_{-t}\|} \sim \frac{e^{-\alpha t}}{e^{-\alpha t \cdot \frac{1}{\rho}}} \tag{9.19}$$

for some $\rho \in \mathbb{R}$ and $\alpha > 0$ as $t \to \infty$ then define for the best asymptotic estimate for $1/\rho$, define $\sigma(p)$ to be

$$\boxed{\sigma(p) = \inf\left\{ b \in \mathbb{R} : \lim_{t\to\infty} \left( \frac{\|w_0\|}{\|w_{-t}\|} \right)^b / \frac{\|v_0\|}{\|v_{-t}\|} = 0,\ \forall w_0 \in N_pM_0,\ \forall v_0 \in T_pM_0 \right\}.} \tag{9.20}$$

The numbers $\nu(p)$ and $\sigma(p)$ are called *F*enichel-type numbers. They are sometimes also referred to as Lyapunov-type numbers or sometimes simply Type-numbers. If $\sigma(p) \sim \frac{1}{\rho} < \frac{1}{s}$ for some $s \in \mathbb{Z}^+$, then the normal decay to $M_0$ is $s$-times stronger than the tangential compression along $M_0$, at least along the trajectory through $p$.

*Remark* 9.2 (Properties of Fenichel-type numbers). First, each are constant along trajectories, and dependent only on the $\alpha$-limit set of the trajectory (the set of limit points as $t \to -\infty$). Also, both $\sigma(p)$ and $\nu(p)$ are upper-semicontinuous in $p$ ($\limsup_{p\to p_0} \sigma(p) \leq \sigma(p_0)$).

to discuss the computation of the Fenichel-type numbers. To this end, define two functions

$$A_t(p) = DF^{-t}(p)\big|_{T_pM_0} :\ T_pM_0 \to T_qM_0;\ q = F^{-t}(p); \tag{9.21}$$

$$B_t(p) = \Pi_p DF^t(F^{-t}(p))\big|_{N_{F^{-t}(p)}M_0} :\ N_{F^{-t}(p)}M_0 \to N_pM_0. \tag{9.22}$$

The reason for the $F^{-t}(p)$ being the inner argument for $B_t$ is that we need to step back in order to safely evaluate the linearized map in forward time. With these functions in hand we can calculate the Fenichel-type numbers equipped with the next theorem.

**Theorem 9.3** (Fenichel (1971))**.** *We have that the following holds*

$$\boxed{\nu(p) = \limsup_{t\to\infty} \|B_t(p)\|^{\frac{1}{t}};} \tag{9.23}$$

*and if* $\nu(p) < 1$*, then*

$$\boxed{\sigma(p) = \limsup_{t\to\infty} \frac{\log\left(\|A_t(p)\|\right)}{-\log\left(\|B_t(p)\|\right)}.} \tag{9.24}$$

*Example* 9.6. Consider the dynamical system

$$\begin{cases} \dot{x} = -x(1-x)(1+x) \\ \dot{y} = -by \end{cases};\quad b > 0. \tag{9.25}$$

Further, define $M_0 = \left\{(x,y) :\ x \in \left[-\frac{3}{2}, \frac{3}{2}\right],\ y = 0\right\}$. Using the previous theorem, it is possible to show

$$\nu(A_0) = \nu(C_0) = e^{-b};\quad \sigma(A_0) = \sigma(C_0) = -\frac{2}{b};\quad \nu(B_0) = e^{-b};\quad \sigma(B_0) = \frac{1}{b}. \tag{9.26}$$

The points $A_0$, $B_0$, and $C_0$ along with the manifold $M_0$ are illustrated in the phase space in Fig. 9.17.

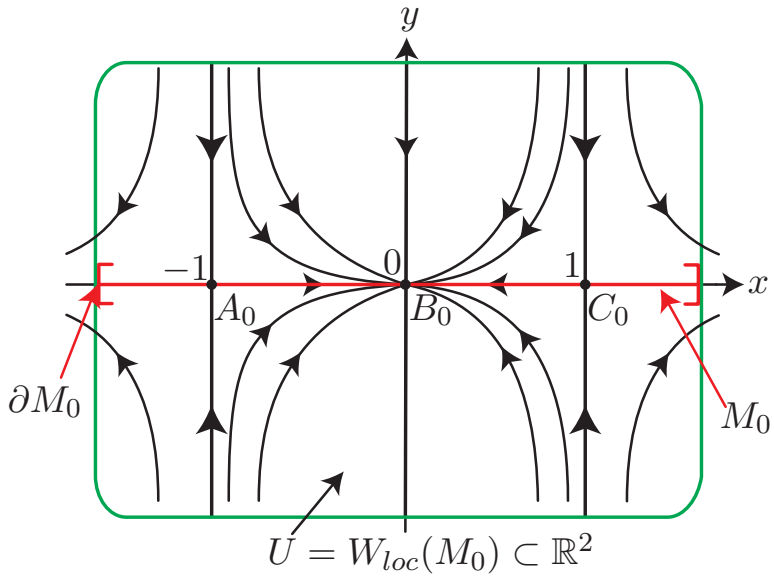


Figure 9.17: $A_0$, $B_0$, and $C_0$ along with the manifold $M_0$ are illustrated in the phase space.

**Definition 9.11.** The manifold $M_0$ is an *$r$-normally hyperbolic invariant manifold* (NHIM) if

(i) $\nu(p) < 1$ for all $p \in M_0$ ;

(ii) $\sigma(p) < \frac{1}{r}$ for all $p \in M_0$.

*Remark* 9.4. If there is no tangential compression, then $\sigma(p) < 0$ and $(ii)$ holds automatically for all $r$.

Building on the previous results, we have the next the next result.

**Theorem 9.5** (Fenichel (1971))**.** *Assume we have*

$$\dot{x} = f_{pert}(x) = f_0(x) + \Delta f(x); \quad f_0, \Delta f \in \mathcal{C}^r, \ r \geq 1, \tag{9.27}$$

*for a compact, overflowing invariant (under $f_0(x)$) manifold $M_0$ which is $\mathcal{C}^r$ and $r$-normally hyperbolic. Then for $\|\Delta f\|_{\mathcal{C}^1}$ small enough, there exists an overflowing invariant manifold $\tilde{M}_0$ for $\dot{x} = f_{pert}(x)$, such that*

*(i) $\tilde{M}_0$ is $\mathcal{C}^1$-close to $M_0$ and $\mathcal{C}^r$-diffeomorphic to $M_0$ ;*

*(ii) All close enough trajectories converge to $\tilde{M}_0$ as $t \to \infty$, as long as they stay in a neighborhood of $\tilde{M}_0$.*

*This result is depicted in Fig. 9.18.*

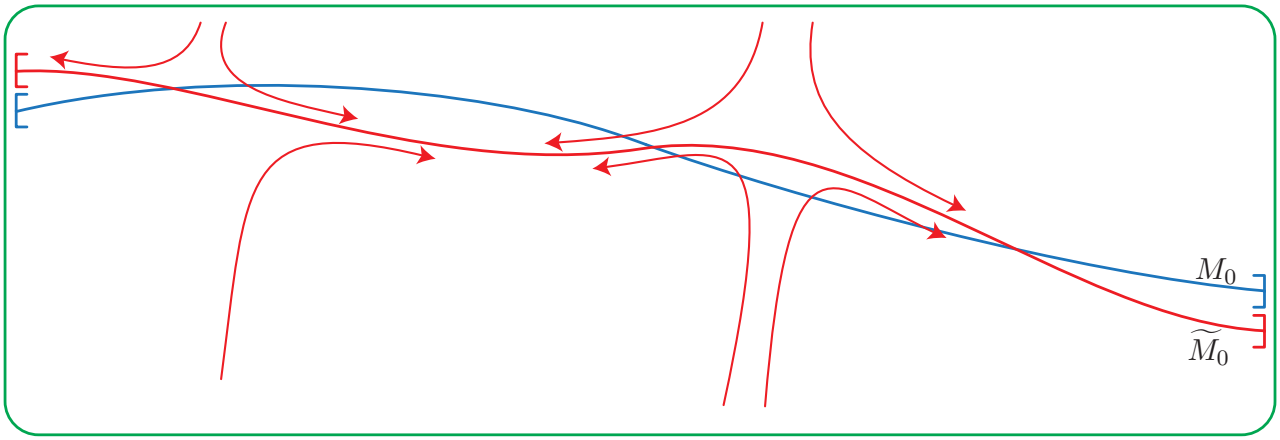


Figure 9.18: The manifolds $M_0$ and $\tilde{M}_0$ illustrated within the locally invariant manifold $W^S_{\text{loc}}(\tilde{M}_0)$ (green).

*Remark* 9.6. The following hold:

(i) If $M_0$ is compact, normally hyperbolic invariant (without boundary), then it attracts all nearby trajectories ($M_0$ is an attractor that smoothly persists);

(ii) The integer $r$ featured in the theorem is the minimum of the integers $r_i \in \mathbb{Z}^+$, where $f \in \mathcal{C}^{r_1}$, $M_0 \in \mathcal{C}^{r_2}$, and $\sigma(p) < \frac{1}{r_3}$;

(iii) Similar results hold for inflowing invariant manifolds that repel nearby initial conditions ($M_0$ and $\tilde{M}_0$ are persistent repellers);

(iv) The theory also extends to saddle-type invariant manifolds too, e.g. the unstable manifold of an overflowing invariant manifold. In this case assume that the tangent bundle $T\mathbb{R}^n|_{M_0}$ has an invariant splitting, i.e. $T_p\mathbb{R}^n = T_pM_0 \oplus N^U_pM_0 \oplus N^S_pM_0$, then there exists an additional Fenichel-type number for $N^UM_0$ and $W^U_{\text{loc}(M_0)}$ is robust. This is shown in Fig. 9.19.

For more see [Wiggins, 1994].

## 9.3 Geometric singular perturbation theory

Singular perturbation theory is the study of systems with two different time scales, i.e. of the form

$$\begin{cases} \dot{x} &= f(x,y,\varepsilon) \\ \varepsilon\dot{y} &= g(x,y,\varepsilon) \end{cases}; \quad 0 \leq \varepsilon \ll 1; \quad f,g \in \mathcal{C}^1. \tag{9.28}$$

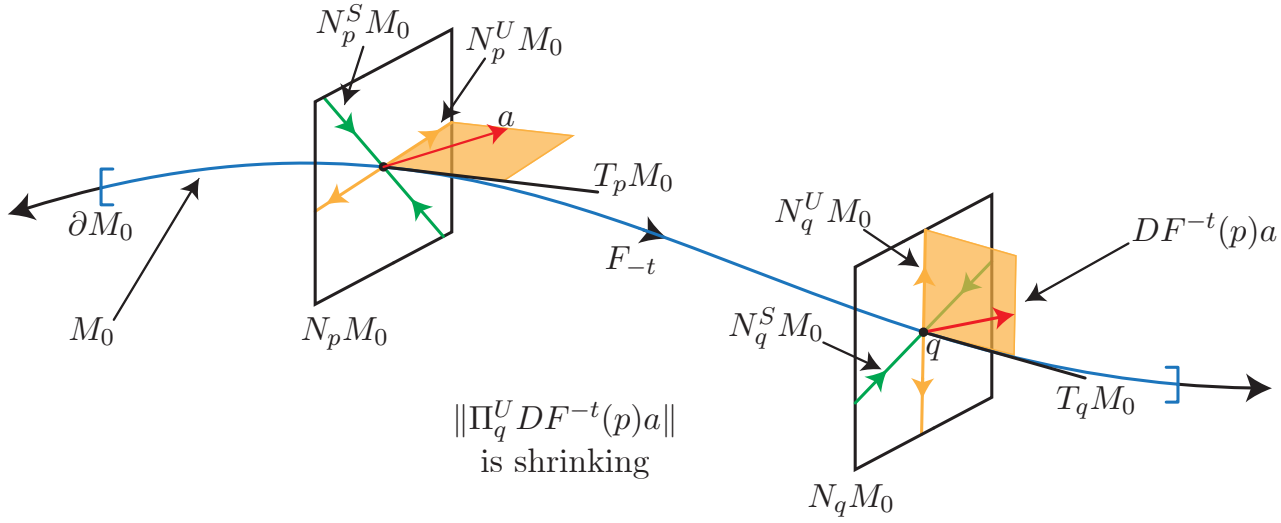


Figure 9.19: The effect of the tangent bundle having an invariant splitting with the addition of a Fenichel-type number for $N^U M_0$ and robust $W^U_{\text{loc}}(M_0)$.

In such systems $x \in \mathbb{R}^n$ is called the *slow variable* and $y \in \mathbb{R}^m$ is the *fast variable*.

*Example* 9.7 (One degree of freedom mechanical oscillator with a small mass). An example of a system with two different time scales is a single degree of freedom mechanical oscillator with a small mass, the dynamical system is given by

$$\varepsilon\ddot{x} + F(x, \dot{x}, \varepsilon) = 0 \implies \begin{cases} \dot{x} &= y \\ \varepsilon\dot{y} &= -F(x, y, \varepsilon) \end{cases}. \tag{9.29}$$

Note that $\dot{y} = \frac{1}{\varepsilon}F(x, y, \varepsilon)$ implies that $y$ changes very quickly. The equations for $\dot{y}$ become singular as $\varepsilon \to 0$, hence we have singular perturbation (the smooth dependence of solutions on $\varepsilon$ does not apply).

The classical approach to such systems is to set $\varepsilon = 0$ in (9.28), obtaining

$$\begin{cases} \dot{x} = f(x, y, 0) \\ 0 = g(x, y, 0) \end{cases}. \tag{9.30}$$

This is a mixed differential-algebraic system of equations. We solve $g(x, y, 0) = 0$ for $y = G_0(x)$, expand in $\varepsilon$, and plug this into $\dot{x} = f(x, y, \varepsilon)$. Sometimes this works, however this is not guaranteed. Instead we will follow [Fenichel, 1979].

First, turn this into a regular perturbation problem by introducing the fast time $t = \varepsilon\tau$, i.e $\tau = \frac{t}{\varepsilon}$ for $0 < \varepsilon \ll 1$. This change of coordinates induces via the chain rule the Jacobian

$$\frac{d(\cdot)}{dt} = \frac{d(\cdot)}{d\tau}\frac{d\tau}{dt}. \tag{9.31}$$

The differential of a function $f$ with respect to $t$ will be denoted by $\dot{f}$, and is equal to $f'\frac{1}{\varepsilon}$, with $f'$ denoting the differential of $f$ with respect to $\tau$. Hence we obtain the rescaled system

$$\begin{cases} x' = \varepsilon f(x, y, \varepsilon) \\ y' = g(x, y, \varepsilon) \end{cases}. \tag{9.32}$$

Note that (9.32) is a regular perturbation problem from a fictitious limit. The systems (9.28) and (9.32) are not equivalent for $\varepsilon = 0$, but are equivalent for any $\varepsilon > 0$.

We analyze the $\varepsilon = 0$ limit next, i.e.

$$\begin{cases} x' = 0 \\ y' = g(x, y, 0) \end{cases}. \tag{9.33}$$

Assume that that there exists a pair $(x_0, y_0)$ such that $g(x_0, y_0, 0) = 0$, i.e. a fixed point of (9.33). Then we have $m$ equations for $m$ unknowns $y_0$, by the Implicit Function Theorem a unique nearby zero $y = \varphi_0(x)$ with $\varphi \in \mathcal{C}^1$ exists whenever we have that the following hold

(i) The matrix $D_y g(x_0, y_0, 0)$ is nonsingular;

(ii) The derivative $D_x g(x_0, y_0, 0)$ exists near $(x_0, y_0)$.

In which case we can define the $n$-dimensional invariant manifold filled with fixed points

$$\boxed{M_0 = \{(x, y) : \ y = \varphi_0(x), \ x \in U, \ U \text{ compact}, \ \det\left[D_y g(x, y, 0)\right] \neq 0\}.} \tag{9.34}$$

This manifold is called the *critical manifold* and an example is shown in Fig. 9.20.

Next, we linearize. Begin by defining $\eta = y - \varphi_0(x)$, then calculate the derivative of $\eta$ with respect to $\tau$

$$\eta' = y' - \underbrace{\varphi_0' x'}_{=0} = g(x, \varphi_0(x) + \eta, 0) = \underbrace{g(x, \varphi_0(x), 0)}_{=0} + D_y g(x, \varphi_0(x), 0)\eta + \mathcal{O}(\eta^2). \tag{9.35}$$

Thus our linearization is $\eta' = D_y g(x, \varphi_0(x), 0)\eta$, due to the nonsingularity condition before we know that $\lambda_i(D_y g(x, y, 0)) \neq 0$ for $i = 1, \ldots, m$. If, in addition, we have that $\mathrm{Re}[\lambda_i(D_y g(x, y, 0)] \neq 0$ for $i = 1, \ldots, m$, then $M_0$ is normally hyperbolic.

The most important case is when $M_0$ is normally attracting, i.e. when $\mathrm{Re}[\lambda_i(D_y g(x, y, 0)] < 0$ for each $i = 1, \ldots, m$. In which case $M_0$ is $r$-normally hyperbolic, where if $\varphi_0(x)$ is $\mathcal{C}^r$ for any $r \in \mathbb{Z}^+$ then $M_0$ persists for $\varepsilon$ small enough.

The classical formal approach often fails when there exist eigenvalues with positive real parts, i.e. when $M_0$ is unstable.

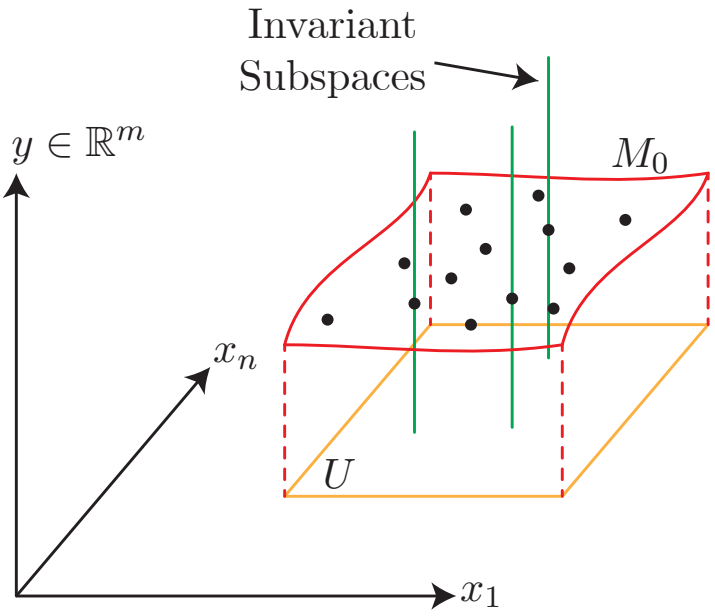


Figure 9.20: An example of the critical manifold $M_0$ (red) filled with fixed points, an $n$-dimensional invariant subspace (green line) passes through each fixed point.

For $\varepsilon > 0$ and with eigenvalues of only strictly negative real parts (as before), we get the perturbation $M_\varepsilon$ of $M_0$. This perturbed manifold is the *slow manifold*, motion on it is of $\mathcal{O}(\varepsilon)$ speed, and in this case it is attracting. We have that for the $\mathcal{C}^1$ function $\varphi_\varepsilon$ (also in $\varepsilon$ as long as $g(x, y, \varepsilon)$ is $\mathcal{C}^1$ in $\varepsilon$)

$$y = \varphi_\varepsilon = \varphi_0(x) + \varepsilon\varphi_1(x) + \mathcal{O}(\varepsilon^2). \tag{9.36}$$

Thus $y$ is a smooth graph over the domain $U$. The perturbed manifold and the representation of $y$ as a smooth graph over $U$ is shown in Fig. 9.21.

We can reduce to the slow manifold and find the reduced flow

$$\begin{aligned} x'|_{M_\varepsilon} &= \varepsilon\, f(x, y, \varepsilon)|_{M_\varepsilon} = \varepsilon f(x, \varphi_\varepsilon(x), \varepsilon) && (9.37)\\ &= \varepsilon\left[f(x, \varphi_0(x), 0) + \varepsilon\left(D_\varepsilon f(x, \varphi_0(x), 0) + D_y f(x, \varphi_0(x), 0)\varphi_1(x)\right)\right] + \mathcal{O}(\varepsilon^3). && (9.38)\end{aligned}$$

In the original time $t$ we find the *reduced flow*

$$\boxed{\dot{x} = f(x, \varphi_0(x), 0) + \varepsilon\left[D_\varepsilon f(x, \varphi_0(x), 0) + D_y f(x, \varphi_0(x), 0)\varphi_1(x)\right] + \mathcal{O}(\varepsilon^2).} \tag{9.39}$$

The leading order restricted flow on the attractor $M_\varepsilon$ is given by $\dot{x} = f(x, \varphi_0(x), 0)$. Any structurally stable feature of this leading-order reduced flow will persist on the attractor. For

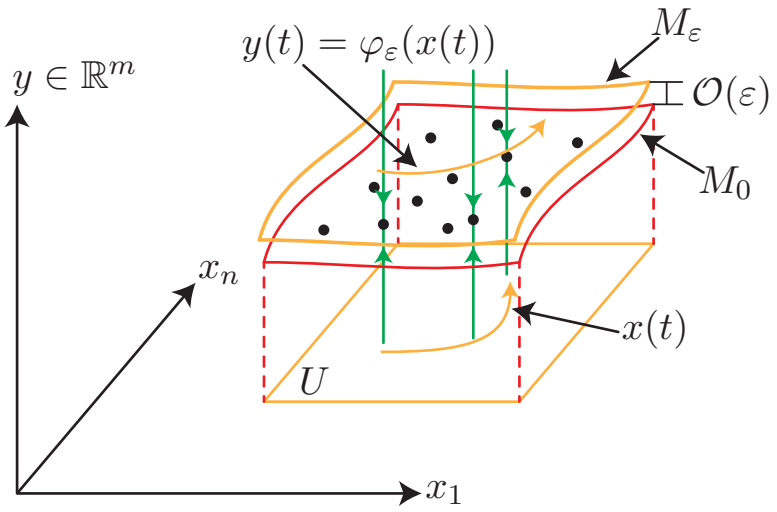


Figure 9.21: The perturbed manifold $M_\varepsilon$ near the critical manifold $M_0$ as well as the representation of $y$ as a smooth graph over the domain $U$.

other features, we need higher-order terms. The dynamics near the slow manifold is attractive towards $M_\varepsilon$ if $\mathrm{Re}[\lambda_i(D_y g(x,y,0)] < 0$ for $i = 1,\dots,m$ for all $(x,y) \in M_0$, i.e. $W^U_{\mathrm{loc}}(M_\varepsilon)$ is empty. Furthermore, we get the stable fibers $f^S_\varepsilon(p)$ which are $\mathcal{C}^1$-close to their respective $f^S_0(p)$, and are of the same dimension $(m)$. These facts are illustrated in Fig. 9.22.

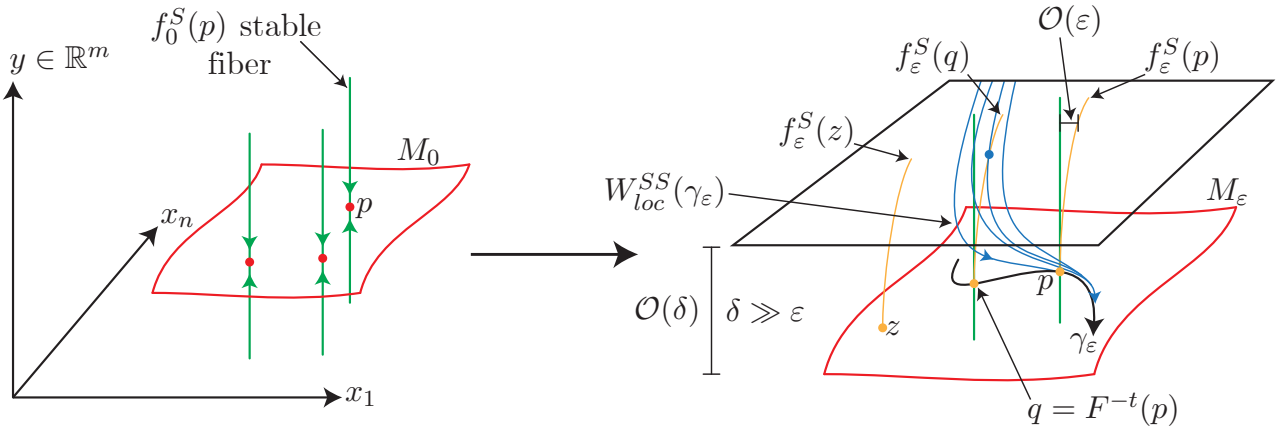


Figure 9.22: The dynamics near the slow manifold and the stable fibers with their relationship to $M_0$ and the stable fibers of the linearization.

Define the $m+1$-dimensional *strong stable manifold* of $\gamma_\varepsilon$ to be

$$W^{SS}_{\mathrm{loc}}(\gamma_\varepsilon) = \bigcup_{p \in \gamma_\varepsilon} f^S_\varepsilon(p), \tag{9.40}$$

i.e. the collection of fastest-converging trajectories to $\gamma_\varepsilon$. By construction, the stable fibers form an invariant family along $\gamma_\varepsilon$, i.e. $F^t(f_\varepsilon^S(q)) \subset f_\varepsilon^S(p)$.

*Example* 9.8 (Clustering of inertial particles in fluids). When finite-size inertial particles are suspended in fluids they tend to form clusters, even in incompressible fluids (as opposed to fluid particles). For a specific case, consider the steady flow in the wake of a bluff body [Burns et al., 1999]. This can be seen in Fig. 9.23.

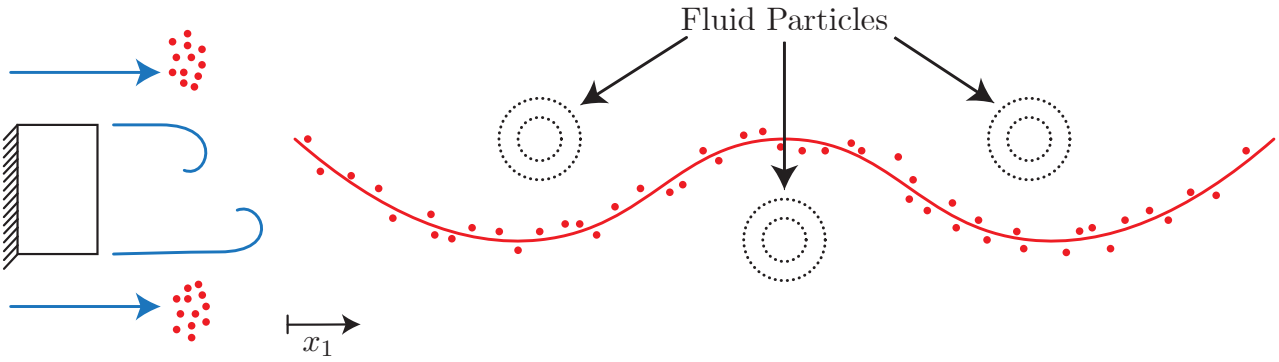


Figure 9.23: The clustering of inertial particles observed experimentally, referred to as the *von Kármán vortex street*.

The notation here will be $x \in \mathcal{S}^1 \times \mathbb{R}$ for the particle position in a moving frame and periodic in the first variable. The steady fluid velocity in the moving frame is given by $u(x) \in \mathbb{R}^2$, the nondimensionalized particle mass is given by $\varepsilon \ll 1$, and the velocity of the inertial particle at time $t$ given by $y = \dot{x} \in \mathbb{R}^2$. The equation of motion in its simplest form is given by [Maxey and Riley, 1983]

$$\begin{cases} \dot{x} &= y \\ \varepsilon \dot{y} &= u(x) - y \end{cases}. \tag{9.41}$$

Thus we have the slow variable $x$ and the fast variable $y$, performing the time transformation as before ($t = \varepsilon\tau$) we obtain

$$\begin{cases} x' = \varepsilon y \\ y' = u(x) - y \end{cases}. \tag{9.42}$$

the $\varepsilon = 0$ limit yields

$$\begin{cases} x' = 0 \\ y' = u(x) - y = g(x, y, 0) \end{cases}. \tag{9.43}$$

Hence, we get the critical manifold

$$M_0 = \left\{(x, y) \in U \subset \mathcal{S}^1 \times \mathbb{R}^3 : \ y = u(x)\right\}. \tag{9.44}$$

To examine the stability of $M_0$ calculate $D_y g(x, u(x), 0) = -I_2$, hence $\mathrm{Re}[\lambda_i] = \lambda_i = -1 < 0$ for $i = 1, 2$. Hence, $M_0$ is normally hyperbolic and attracting, the unperturbed fibers $f_0^S(p)$ for $p \in M_0$ are 2-dimensional planes. The unperturbed geometry is depicted in Fig. 9.24.

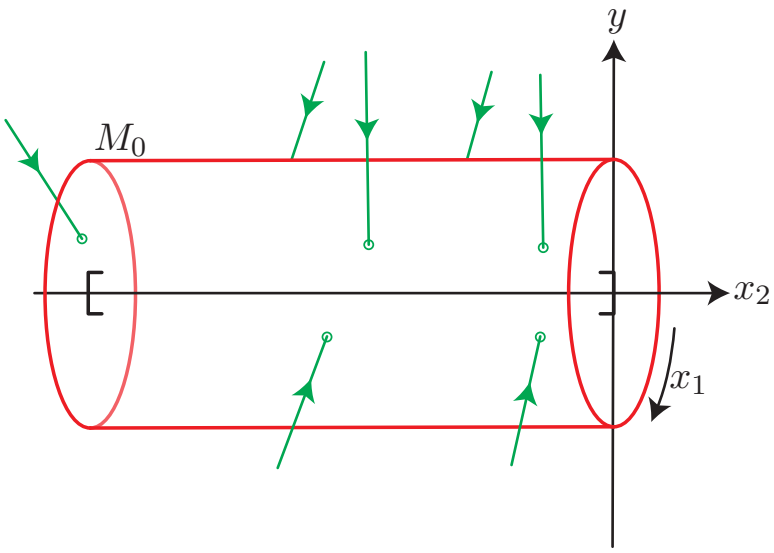


Figure 9.24: The unperturbed geometry of the dynamics given in the example of inertial particles suspended in the fluid.

For $\varepsilon > 0$ we get the perturbed manifold $M_\varepsilon$ which is $\mathcal{C}^1$ $\mathcal{O}(\varepsilon)$-close to $M_0$. Trajectories in $W^{SS}(\gamma_\varepsilon)$ quickly synchronize with $\gamma_\varepsilon$. The perturbed geometry is shown in Fig. 9.25.

To study the dynamics on the slow manifold, write the perturbed manifold $M_\varepsilon$ as the graph of the function $\varphi_\varepsilon$

$$M_\varepsilon = \left\{ (x, y) : \ y = \varphi_\varepsilon(x) = \underbrace{\varphi_0(x)}_{u(x)} + \varepsilon h_1(x) + \mathcal{O}(\varepsilon^2) \right\}. \tag{9.45}$$

By the invariance of $M_\varepsilon$ the formula $y = \varphi_\varepsilon(x)$ holds along full trajectories. Plugging this in and differentiating with respect to $\tau$ we find

$$y' = (Du)x' + \varepsilon(Dh_1)x' + \ldots = (Du)\varepsilon u(x) + \mathcal{O}(\varepsilon^2). \tag{9.46}$$

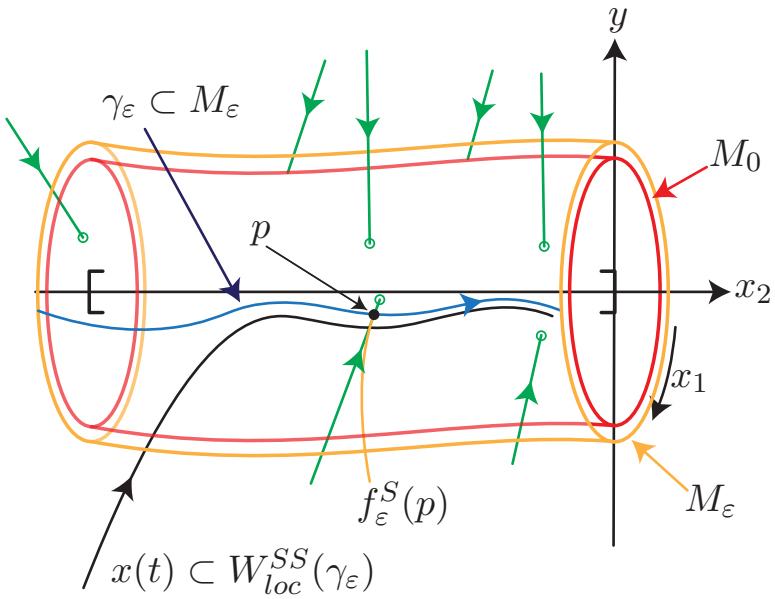


Figure 9.25: The perturbed geometry of the dynamics given in the example of inertial particles suspended in the fluid.

In the second equation the fact $x' = \varepsilon\varphi_\varepsilon(x)$ was used. Also from the equation of motion we find

$$y'|_{M_\varepsilon} = (u(x) - y)|_{M_\varepsilon} = u(x) - (u(x) + \varepsilon h_1(x) + \ldots) = -\varepsilon h_1(x) + \mathcal{O}(\varepsilon^2). \tag{9.47}$$

By comparing coefficients we then find

$$h_1 = -(Du)u. \tag{9.48}$$

Hence, the reduced flow on the slow manifold is

$$x' = \varepsilon\varphi_\varepsilon(x) = \varepsilon\left(u(x) - \varepsilon(Du(x))u(x)\right) + \mathcal{O}(\varepsilon^2). \tag{9.49}$$

Transforming this back into the original time we find

$$\dot{x} = u(x) - \varepsilon(Du(x))u(x) + \mathcal{O}(\varepsilon^2). \tag{9.50}$$

The $\mathcal{O}(1)$ term, $u(x)$, denotes the incompressible terms, meanwhile the $\mathcal{O}(\varepsilon)$ term is the dissipative terms (i.e. has nonzero divergence). This nonzero divergence can be calculated

$$\nabla \cdot [(Du)u)] = \mathrm{Tr}\,[DuDu] = -2Q; \quad Q = \frac{1}{2}\left|\|\Omega\|^2 - \|S\|^2\right|, \tag{9.51}$$

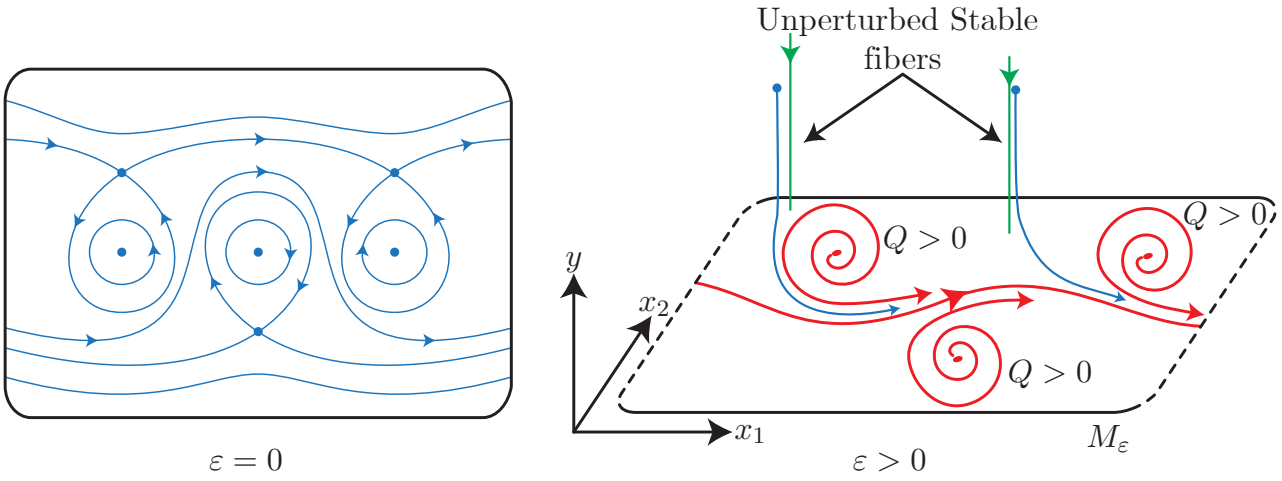


Figure 9.26: Left: The unperturbed phase space. Right: The perturbed phase space for $Q > 0$, the green arrows denotes unperturbed stable fibers, and the red line designates the limit cycle. By identifying the left and right sides, we get a periodic variable.

where $\|\cdot\|$ denotes the Euclidean matrix norm. The unperturbed system and $Q > 0$ case are depicted in Fig. 9.26.

A more general approach to particle motion is given by [Haller and Sapsis, 2008] using the Maxey-Riley equations. For neutrally buoyant particles (ones which are suspended in the fluid) we have

$$\varepsilon \dot{v} - \varepsilon \frac{Du}{Dt} = -(v - y); \quad \varepsilon = \frac{3St}{2} \ll 1, \tag{9.52}$$

where $S$ denotes the Strouhal number. Formulating this in terms of geometric singular perturbation theory, with $x \in \mathbb{R}^d$ $(d = 2, 3)$ we get

$$\begin{cases} \dot{x} = y \\ \varepsilon \dot{y} = u(x,t) - y + \varepsilon \dfrac{Du(x,t)}{Dt} \end{cases} \implies \begin{cases} t' = \varepsilon \\ x' = \varepsilon y \\ y' = u(x,t) - y + \varepsilon \dfrac{Du(x,t)}{Dt} \end{cases}. \tag{9.53}$$

The time transformation used was $t = t_0 + \varepsilon\tau$. From here we have the slow variables $x$ and $t$, and the fast variable $y$. There then exists a $(d+1)$-dimensional slow manifold $M_\varepsilon$ where we have $y = u(x,t) + \varepsilon u_1(x,t) + \mathcal{O}(\varepsilon)$. The reduced particle dynamics here is given by

$$y' = \nabla u x' + u_t \varepsilon + \varepsilon \nabla u_1 x' + u_{1t}\varepsilon^2 + \mathcal{O}(\varepsilon^2) = \varepsilon(\nabla u y + u_t) + \mathcal{O}(\varepsilon^2) \tag{9.54}$$

$$= \varepsilon(\nabla u u + u_t) + \mathcal{O}(\varepsilon^2) = \varepsilon \frac{Du}{Dt} + \mathcal{O}(\varepsilon^2). \tag{9.55}$$

Also, we have

$$y' = u - \left(u + \varepsilon u_1 + \mathcal{O}(\varepsilon^2)\right) + \varepsilon\left(u_t + u_x u\right) \tag{9.56}$$

$$= -u_1 + \varepsilon \frac{Du}{Dt} + \mathcal{O}(\varepsilon^2). \tag{9.57}$$

By comparing terms we see that $u_1 = 0$, in fact this holds up to any order.

The reduced dynamics on the slow manifold follow $\dot{x} = v = u(x,t)$ up to any order independent of $\varepsilon$. This suggests a fast synchronization with the fluid, yet scattering of the particles is still observed along fluid motions. The reason for this can be seen by rewriting (9.52)

$$\dot{v} - u_t - \nabla u u = \mu(u - v); \quad \dot{v} - u_t - \nabla u v = \mu(u - v) + \nabla u(u - v). \tag{9.58}$$

This implies that

$$\begin{cases} \dfrac{d}{dt}\left(v - u(x,t)\right) = -\left(\nabla u + \mu I\right)(v - u) \\ \dfrac{d}{dt} x = v \end{cases}. \tag{9.59}$$

These in turn imply that $v = u(x,y)$ is an invariant manifold for any $\varepsilon = \frac{1}{\mu}$, but we only know its stability for $\varepsilon \ll 1$. This means that different parts of the manifold may have different stability types.

To analyze the stability in more detail set $z = v - u(x,y)$, i.e. the distance from $M_\varepsilon$ in the $y$-direction. The dynamics with this variable now is

$$\begin{cases} \dot{z} = -\left[\nabla u(x,t) + \mu I\right] z \\ \dot{x} = z + u(x,t) \end{cases} \implies \frac{1}{2}\frac{d}{dt}\|z\|^2 = -\langle z, (\nabla u + \mu I) z\rangle = -\langle z, (S + \mu I) z\rangle, \tag{9.60}$$

Where $S = \frac{1}{2}\left(\nabla u + \nabla u^T\right) = S^T$ is the rate of strain tensor.

We can derive conditions now for guaranteed attraction or repulsion:

(i) **Exponential attraction** is guaranteed when we have

$$\frac{1}{2}\frac{d}{dt}\|z\|^2 \leq \lambda_{\max}\left[-(S + \mu I)\right]\|z\|^2. \tag{9.61}$$

By the Gronwall-inequality, this implies

$$\|z(t)\| \leq \|z(t_0)\| \exp\left(\int_{t_0}^{t} \lambda_{\max}[-(S + \mu I)] ds\right). \tag{9.62}$$

Thus we have that $\|z(t)\|$ decays as long as we have

$$\lambda_{\max}\left[-S(x(t), y(t), t) - \mu I\right] < 0 \Leftrightarrow \lambda_{\min}\left[S(x,t) + \mu I\right] > 0 \Leftrightarrow \mu > \sqrt{|\det(S)|}. \tag{9.63}$$

(ii) **Exponential repulsion** is guaranteed when we have

$$\frac{1}{2}\frac{d}{dt}\|z\|^2 \geq \lambda_{\min}\left[-(S+\mu I)\right]\|z\|^2. \tag{9.64}$$

Thus we have that $\|z(t)\|$ grows instantaneously when

$$\lambda_{\max}\left[-(S+\mu I)\right] > 0 \implies \lambda_{\max}\left[S+\mu I\right] < 0. \tag{9.65}$$

More generally, an attracting/repelling invariant manifold may have localized repulsion/attraction along it, even if all trajectories approach it asymptotically [Sapsis and Haller, 2010].

To study the local repulsion/attraction, consider the following dynamical system

$$\dot{x} = f(x,t); \quad x \in \mathbb{R}^n, \tag{9.66}$$

with the flow map $F_{t_0}^t : x_0 \mapsto x(t;t_0,x_0)$ and a locally invariant $k$-dimensional manifold $\mathcal{M}(t) \subset \mathbb{R}^n$, i.e. $F_{t_0}^t(\mathcal{M}(t_0)) \subset \mathcal{M}(t)$. Now define the *normal infinitesimal Lyapunov exponent* $\sigma(p,t)$ (abbreviated NILE) by taking the $t \to t_0$ limit and looking at the largest growth for a tangent vector $w_0$, i.e.

$$\boxed{\sigma(p,t) = \lim_{s\to 0}\frac{1}{s}\log\left\|\Pi^{t+s}_{F_t^{t+s}(p)} DF_t^{t+s}(p)\Big|_{N_p\mathcal{M}(t)}\right\|.} \tag{9.67}$$

The graphical depiction of the terms used for the calculation of the normal infinitesimal Lyapunov exponent are depicted in Fig. 9.27.

To compute the NILE we can use the following formula

$$\boxed{\sigma(p,t) = \max_{n\in N_p\mathcal{M}(t),\ |n|=1}\langle n, S(p,t)n\rangle; \quad S = \frac{1}{2}\left(Df + Df^T\right).} \tag{9.68}$$

Let us express the invariant manifold as a graph over $k$ variables, i.e. write $x$ as $x = (a,b)$, where $a \in \mathbb{R}^k$ and $b \in \mathbb{R}^{k-n}$, $f = (A,B)$ and on the invariant manifold $M$ we have $b = h(a)$. We can then write

$$\sigma(a,t) = \frac{1}{2}\lambda_{\max}\left[\Gamma(a,t) + \Gamma(a,t)^T\right]; \tag{9.69}$$

where we have defined the tensors

$$\Gamma(a,t) = B_b(a,h(a),t) - Dh(a,t)A_b(a,h(a),t). \tag{9.70}$$

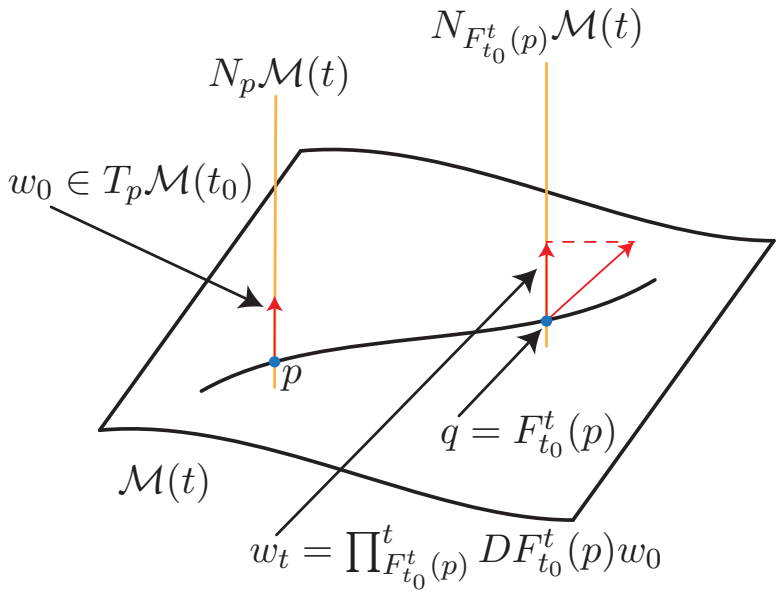


Figure 9.27: The geometry of tangent vectors along the invariant manifold M. The quantities used to define the NILE (9.67) are indicated in the figure.

*Example* 9.9 (Neutrally buoyant particles). Consider the dynamical system with $x, z \in \mathbb{R}^2$

$$\begin{cases} \dot{x} = z + u(x,t) \\ \dot{z} = -\left(\nabla u(x,t) + \mu I\right) z \end{cases}; \implies \mathcal{M}_0(t) \left\{(x,z) \in \mathbb{R}^2 \times \mathbb{R}^2 : \ z = 0\right\}. \tag{9.71}$$

Therefore we have

$$\Gamma(a,t) = B_b(x,0,t) = -\left(\nabla u(x,t) + \mu I\right). \tag{9.72}$$

With this we can calculate the NILE

$$\sigma(p,t) = \lambda_{\max}\left[-(S(x,t) + \mu I)\right]; \quad S = \frac{\nabla u + \nabla u^t}{2}. \tag{9.73}$$

As concluded before, where $\sigma(p,t) > 0$ we have a region of local repulsion, analogously where $\sigma(p,t) < 0$ we have a region of local attraction.

## 9.4 Existence of invariant manifolds

We now turn to the various methods to prove the existence of an invariant manifold:

(i) **Lyapunov-Perron Method** The manifold is the solution of a global integral equation, this approach requires global coordinates;

(ii) **Hadamard's Method** The manifold is sought as an invariant graph, this was the approach Fenichel used and local coordinates are sufficient for this approach;

(iii) **Topological Methods** There are multiple topological approaches, for instance the Wasewsky principle.

Next, the Lyapunov-Perron method will be sketched for stable manifolds on the plane. Consider the dynamical system

$$\dot{x} = f(x); \quad x \in \mathbb{R}^2; \quad f(0) = 0; \quad \text{spect}[Df(0)] = \{-\lambda_S, \lambda_U\}; \quad \lambda_S, \lambda_U > 0. \tag{9.74}$$

The objective is to construct a stable manifold tangent to $E_S$. This will be achieved in multiple steps.

(i) First, choose a good basis

$$x = Ty; \quad T = \begin{pmatrix} e_S & e_U \end{pmatrix}; \quad y = \begin{pmatrix} y_S \\ y_U \end{pmatrix}. \tag{9.75}$$

And implement this transformation

$$\dot{x} = Df(0)x + \mathcal{O}(\|x\|^2) \implies T\dot{y} = Df(0)Ty + \mathcal{O}(\|y\|^2). \tag{9.76}$$

By multiplying from the left by $T^{-1}$ we find

$$\dot{y} = \underbrace{T^{-1}DF(0)T}_{\text{diag}[-\lambda_S, \lambda_U]} y + \mathcal{O}(\|y\|^2); \quad \lim_{y \to 0} \frac{\|g(y)\|}{\|y\|^2} < \infty. \tag{9.77}$$

And now we find

$$\begin{cases} \dot{y}_S = -\lambda_S y_S + g_S(y) \\ \dot{y}_U = \lambda_U y_U + g_U(y) \end{cases}. \tag{9.78}$$

(ii) Second, define the local stable manifold. The classic definition of $W^S(0)$ is that it is the collection of initial coordinates $y_0$ which under the limit of the flow map have $\lim_{t \to \infty} F_t(y_0) = 0$. This requires detailed knowledge of asymptotics, and thus is hard to work with. For linear systems an alternative definition is, that $W^S(0)$ is the collection of forward-bounded trajectories, but this is generally nothing special after the addition of nonlinear terms. However, we may change the nonlinear system away from the origin to ensure that boundedness remains a distinguishing property.

To go about this we separate the phase space into the rectangular area (of half-width $\delta$) around the origin and the rectangular annulus of outer half-width $2\delta$. In the inner rectangle, the original nonlinear dynamics hold, outside of the annulus the dynamics are 0 (i.e. $f(x) = 0$). Interpolating in between these regimes takes place on the annulus using a so-called *bump function* $\Psi_\delta$, a smooth function which is equal to 1 on the interval $[-\delta, \delta]$ and equal to 0 on the intervals $(-\infty, -2\delta)$ and $(2\delta, \infty)$, in between these regions is the interpolation. The bump function and the interpolated new dynamical system are depicted in Fig. 9.28.

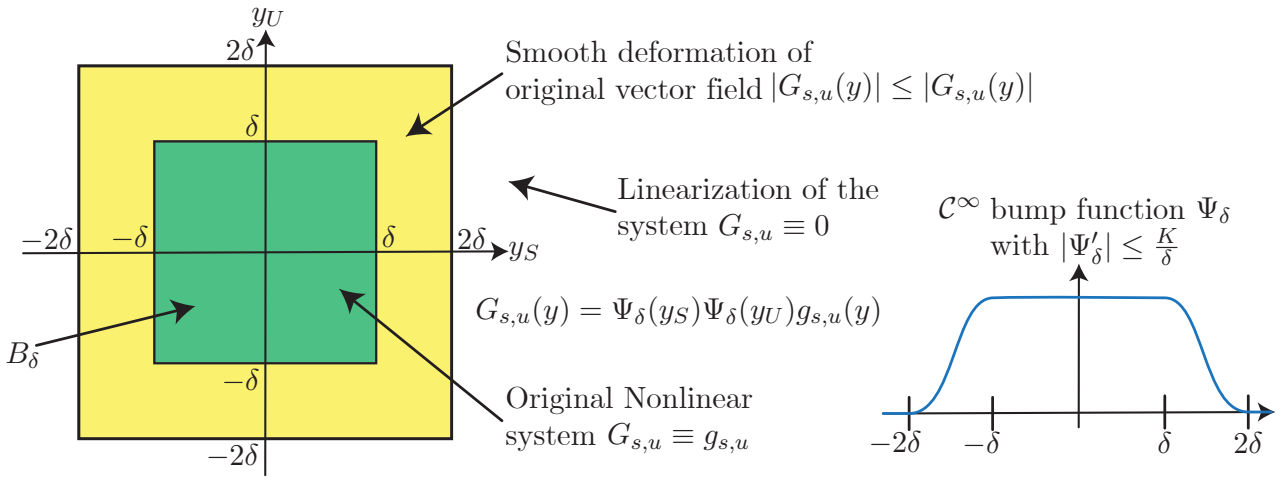


Figure 9.28: The interpolated dynamical system, within the green area the original nonlinear dynamics hold, on the yellow area the smoothly deformed (interpolated) system holds, and outside of these two regions, the null dynamics hold. The bump function is also depicted.

Thus we have the modified system

$$\begin{cases} \dot{y}_S = -\lambda_S y_S + G_S(y) \\ \dot{y}_U = \lambda_U y_U + G_U(y) \end{cases}. \tag{9.79}$$

Furthermore, we can bound this system

$$|G'_{S,U}(y)| = \Psi'_\delta g_{S,U} + \Psi_\delta g'_{S,U} \leq |\Psi_\delta||g_{S,U}| + |\Psi_\delta||g'_{S,U}| \tag{9.80a}$$

$$\leq \frac{K}{\delta} C\delta^2 + 1 \cdot 2\delta \leq C_0 \delta. \tag{9.80b}$$

Then our goal becomes to find globally bounded solutions in the modified system, i.e.

$$W^S_{\text{loc}}(0) = \left\{ y_0 \in B_\delta : \sup_{t \geq 0} \left| \mathcal{F}^t(y_0) \right| < \infty \right\}. \tag{9.81}$$

(iii) Now pretend that $y(t)$ is known to find the *equivalent integral equations*. The variation of constants formula yields

$$y_S(t) = y_S(t_S)e^{-\lambda_S(t-t_S)} + \int_{t_S}^{t} e^{-\lambda_S(t-s)}G_S(y(s))ds \tag{9.82a}$$

$$y_U(t) = y_U(t_U)e^{\lambda_U(t-t_U)} + \int_{t_U}^{t} e^{\lambda_U(t-s)}G_U(y(s))ds. \tag{9.82b}$$

By taking $t_U \to \infty$ and $t_S = 0$ we restrict to $W^S_{\text{loc}}(0)$, thus we find the integral equations for $W^S_{\text{loc}}(0)$

$$\begin{cases} y_S(t) = e^{-\lambda_s t}y_S(0) + \int_0^t e^{-\lambda_S(t-s)}G_S(y(s))ds \\ y_U(t) = \int_\infty^t e^{\lambda_U(t-s)}G_U(y(s))ds \end{cases}. \tag{9.83}$$

(iv) Finally we have a fixed point problem, the equation (9.83) is fulfilled when $y(t) = \mathcal{F}(y(t); y_S)$, thus we have a fixed point problem with the initial condition as a parameter. There exists a unique solution if $\mathcal{F}$ is a contraction mapping from a complete metric space into itself, i.e. when we can apply the Banach fixed point theorem. For $\mathcal{F}$ to be contracting, we require

$$d(\mathcal{F}(x), \mathcal{F}(y)) \leq qd(x, y); \quad 0 < q < 1, \tag{9.84}$$

for the metric $d$ defined over the underlying metric space.

Define the set

$$\mathcal{B}_\delta = \left\{ y(t) : [0, \infty) \to \mathbb{R}^2 : \ y_S(0) = \delta, \ \|y\| \leq K, \ y \in \mathcal{C}^0[0, \infty) \right\}, \tag{9.85}$$

where the norm is $\|y\| = \sup_{t\geq 0} \|y(t)\|$. The space can be equipped with the metric

$$d(x, y) = \sup_{t\geq 0} \|x(t) - y(t)\| = \|y - x\|, \tag{9.86}$$

and is complete with respect to this metric by the uniform convergence of Cauchy sequences of $\mathcal{C}^0$ functions to $\mathcal{C}^0$ functions, with limits obeying the same bound. We also have to show that $\mathcal{F}$ maps into itself, i.e. that $\|\mathcal{F}(y)\| \leq K$ when $\|y\| \leq K$. To this end,

calculate

$$\begin{aligned}
\|y(t)\| &\leq |y_S(t)| + |y_U(t)| && \text{(9.87a)}\\
&\leq e^{-\lambda_S(t)}|y_S(0)| + \int_0^t e^{-\lambda_S(t-s)}|G_S(y(s))|ds + \int_t^\infty e^{\lambda_U(t-s)}|G_U(y(s))|ds && \text{(9.87b)}\\
&\leq \delta + e^{-\lambda_S t}\int_0^t e^{\lambda_S s}\underbrace{|G_S(y(s))|}_{\leq C\delta^2}ds + e^{\lambda_U t}\int_t^\infty e^{-\lambda_U s}\underbrace{|G_U(y(s))|}_{\leq C\delta^2}ds && \text{(9.87c)}\\
&\leq \delta + C\delta^2\frac{1}{\lambda_S}\left[e^{-\lambda_S(t-s)}\right]_0^t - \frac{C\delta^2}{\lambda_U}\left[e^{\lambda_U(t-s)}\right]_t^\infty && \text{(9.87d)}\\
&= \delta + C\delta^2\left(\frac{1}{\lambda_S}\left(1-e^{-\lambda_S t}\right) + \frac{1}{\lambda_U}(1-0)\right) && \text{(9.87e)}\\
&\leq \delta + C\delta^2\left(\frac{1}{\lambda_S} + \frac{1}{\lambda_U}\right) \leq K. && \text{(9.87f)}
\end{aligned}$$

This shows that $y(t) \in \mathcal{B}_\delta$ for $\delta$ small enough. Furthermore, we have that $\mathcal{F}$ is a contraction mapping for $\delta$ small enough.

$$\begin{aligned}
\|\mathcal{F}(y) - \mathcal{F}(x)\| &\leq \int_0^t e^{-\lambda_S(t-s)}|G_S(x(s)) - G_S(y(s))|ds && \text{(9.88)}\\
&\quad + \int_t^\infty e^{\lambda_U(t-s)}|G_U(x(s)) - G_U(y(s))|ds && \text{(9.89)}\\
&\leq \int_0^t e^{-\lambda_S(t-s)}C_0\delta\|x(s) - y(s)\|ds && \text{(9.90)}\\
&\quad + \int_t^\infty e^{\lambda_U(t-s)}C_0\delta\|x(s) - y(s)\|ds && \text{(9.91)}\\
&\leq C_0\delta\left(\frac{1}{\lambda_S} + \frac{1}{\lambda_U}\right)\|x - y\|, && \text{(9.92)}
\end{aligned}$$

where in (9.91) the mean value theorem was used. This inequality implies that for $\delta$ small enough,

$$\|\mathcal{F}(y) - \mathcal{F}(x)\| \leq \underbrace{C_0\delta\left(\frac{1}{\lambda_S} + \frac{1}{\lambda_U}\right)}_{\leq q<1}\|x - y\|. \tag{9.93}$$

For the general, mulit-dimensional case see [Chicone, 1999]. A similar argument also work for the strong stable manifold.

## 9.5 Spectral submanifolds and model reduction

For reference, see [Haller and Ponsioen, 2016]. A typical model reduction involves a projection onto subspaces, even for nonlinear systems.

*Example* 9.10. consider the dynamical system

$$\begin{cases} \dot{x} = xy + x^3 \\ \dot{y} = -y - 2x^2 \end{cases}. \tag{9.94}$$

Next, linearize at $x = y = 0$ to find a stable and center manifold as shown in Fig. 9.29. The

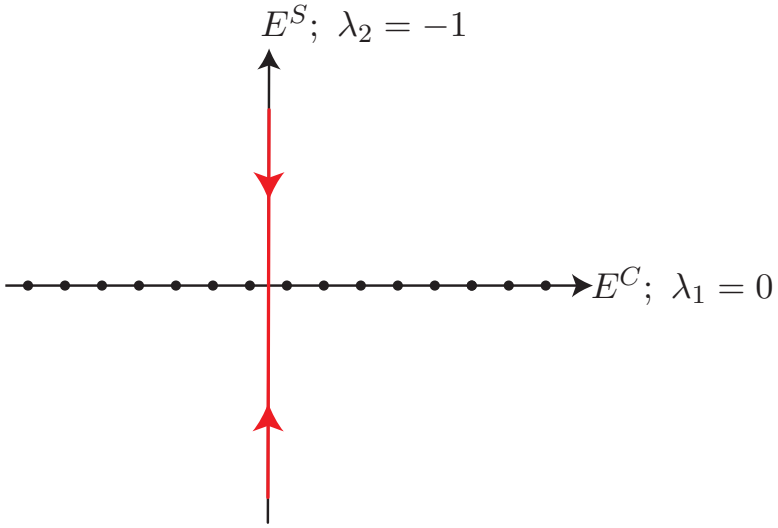


Figure 9.29: Linearization of the dynamical system (9.94)

usual argument is that fast-decaying modes die out. Thus we project the system onto the center subspace $E^C$, which is equivalent to setting $y = 0$ in (9.94) to find

$$\dot{x} = x^3. \tag{9.95}$$

This is the *reduced-order model from projection*. Here the origin is unstable, and all non-equilibrium orbits are unbounded.

The dynamical systems approach is an exact model reduction. We known that a center manifold $W^C(0)$ tangent to the center subspace at 0 exists and is $\mathcal{C}^r$ for any $r \geq 0$, but is not unique in general. The best possible reduced-order model examines the dynamics on $W^C(0)$. Write

$$W^C(0) = \{(x, y) :\ y = h(x) = ax^2 + bx^3 + cx^4 + \mathcal{O}(x^5)\}. \tag{9.96}$$

By invariance we know that

$$\dot{y} = h'\dot{x} = \left[2ax + \mathcal{O}(x^2)\right]\left(xy + x^3\right) = 2a(a+1)x^4 + \mathcal{O}(x^5). \tag{9.97}$$

By restricting to $W^C(0)$ we also have that

$$\dot{y}|_{W^C(0)} = -y|_{W^C(0)} - 2x^2 = -(a+2)x^2 - bx^3 - cx^4 + \mathcal{O}(x^5). \tag{9.98}$$

Comparing coefficients allows us to obtain $a = -2$, $b = 0$, and $c = -2a(a+1) = -4$. Therefore we find the reduced dynamics on $W^C(0)$

$$\dot{x} = (xy + x^3)\big|_{W^C(0)} = x(-2x^2 + \mathcal{O}(x^4)) + x^3 = -x^3 + \mathcal{O}(x^5). \tag{9.99}$$

This is the exact reduced-order model and we find that $x$ is asymptotically stable. The reduced-order model captures the correct local dynamics. Next, we discuss if this discrepancy is the result of a vanishing eigenvalue, i.e. a degenerate linear part.

*Example* 9.11. A more realistic case is when all modes decay under linearization, i.e. $W^C(0) = W^U(0) = \emptyset$. Consider the dynamical system

$$\begin{cases} \dot{x} = -\frac{1}{4}x + xy + x^3 \\ \dot{y} = -y - 2x^2 \end{cases}. \tag{9.100}$$

Once again we linearize at the origin and find that $W^S(0) = \mathbb{R}^2$. The usual argument here would be to project to the slowest-decaying mode, this implies projecting onto the $x$-axis by setting $y = 0$. We find the reduced order model

$$\dot{x} = x\left(-\frac{1}{4} + x^2\right). \tag{9.101}$$

The origin is asymptotically stable with this reduction, whose purpose was to understand the nonlinear dynamics. The nonlinear model prediciton is to have two nontrivial unstable fixed points and a domain of attraction. We are left hoping that this is correct, because there is no way to refine this projection further, as every mode except one is already included.

The fixed points of the full system satisfy

$$\begin{cases} 0 = -y - 2x^2 \\ 0 = -\frac{1}{4}x + xy + x^3 \end{cases} \implies -\frac{1}{4}x - x^3 = 0 \Leftrightarrow -x\left(\frac{1}{4} + x^2\right) = 0. \tag{9.102}$$

Hence, no fixed point exists! The projection onto subspaces can give incorrect results even if all but one mode in included in the reduced model. We ask if it is possible to do exact

model reduction onto slowly decaying modes. The issues of nonuniqueness in linear systems and smoothness of weak stable manifolds present themselves. Consider

$$\begin{cases} \dot{x} = \lambda_1 x \\ \dot{y} = -\lambda_2 y \end{cases} \implies \begin{cases} x(t) = e^{-\lambda_1 t} x_0 \\ y(t) = e^{-\lambda_2 t} y_0 \end{cases}. \tag{9.103}$$

We can then calculate the equations for the manifolds

$$\left(\frac{x}{x_0}\right)^{\lambda_1} = \left(\frac{y}{y_0}\right)^{\lambda_2} \implies y = \frac{y_0}{x_0^{\frac{\lambda_1}{\lambda_2}}} x^{\frac{\lambda_1}{\lambda_2}} = C_0 x^{\frac{\lambda_1}{\lambda_2}}. \tag{9.104}$$

See that due to $C_0$ being arbitrary we have nonuniqueness. Also, the slow manifold is only $\lfloor \frac{\lambda_1}{\lambda_2} \rfloor$-times differentiable unless a resonance arises ($\lambda_1 = k\lambda_2$). But, there is a unique $\mathcal{C}^\infty$, even analytic slow manifold with $C_0 = 0$, i.e. $y = 0$. The smoothest slow manifold is unique. The question remains if this also happens in nonlinear systems.

Consider the setup

$$\dot{x} = Ax + f_0(x); \quad x \in \mathbb{R}^n; \quad f_0(x) = \mathcal{O}(\|x\|^2); \quad f_0 \in \mathcal{C}^r, \ r \in \mathbb{N}^+ \cup \{\infty, a\}. \tag{9.105}$$

If $r = a$, then $f_0$ is analytic. From the definition of $f_0$ we can see that $x = 0$ is a fixed point. The eigenvalues of $A$ are $\lambda_1, \ldots, \lambda_N$ sorted by descending real part, i.e. $\mathrm{Re}[\lambda_N] \leq \ldots \leq \mathrm{Re}[\lambda_1] < 0$. Therefore $x = 0$ is asymptotically stable. The eigenspaces are given by $E_j$ and are the span of the real and imaginary parts of all eigenvectors and generalized eigenvectors, corresponding to $\lambda_j$. Hence we have $\mathbb{R}^n = \bigoplus_{j=1}^N E_j$.

We call the direct sum of an arbitrary number of distinct eigenspaces the *spectral subspace*

$$E = \bigoplus_{j=1}^{t} E_j; \quad 1 \leq k \leq N, \tag{9.106}$$

this is always an invariant subspace for $\dot{x} = Ax$. The *spectral quotient of* $A$ is given by

$$\sigma(E) = \left\lfloor \frac{\text{smallest } \mathrm{Re}[\lambda_k] \text{ outside of } E}{\text{largest } \mathrm{Re}[\lambda_i] \text{ inside of } E} \right\rfloor. \tag{9.107}$$

An illustration of these values are given in Fig. 9.30.

We can check in a 2-dimensional linear example, that this spectral quotient was the degree of differentiability for the general slow manifolds.

The *external nonresonance condition for* $E$ for $\dim[E] = q$ is

$$\langle m, \lambda \rangle_E = m_1\lambda_{j1} + \ldots + m_1\lambda_{jq} \neq \lambda_l, \ m_i \in \mathbb{N}, \tag{9.108}$$

where $\lambda_{j1}, \ldots, \lambda_{jq} \in \mathrm{spect}[A|_E]$ and $\lambda_l \notin \mathrm{spect}[A|_E]$.

Now we ask if the full nonlinear system has a *spectral submanifold* (SSM), i.e. a nonlinear continuation of $E$.

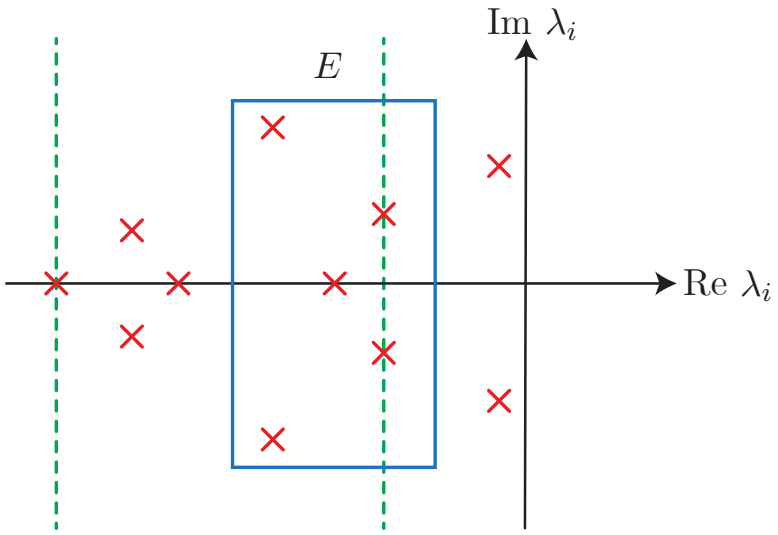


Figure 9.30: The eigenvalue constellation, the left dotted green line designates the smallest real part outside of $E$ and the right dotted green line designates the largest real part inside of $E$.

**Theorem 9.7.** *Assume the low-order external nonresonance condition*

*(i)* $\langle m, \lambda \rangle_E \neq \lambda_l$ *for* $\lambda_l \neq spect[A|_E]$ ;

*(ii)* $2 \leq |m| \leq \sigma(E)$ *with the norm* $|m| = m_1 + \ldots + m_1$;

*then we have that the following hold:*

*(i) There exists a* $\mathcal{C}^r$ *spectral submanifold* $W(0)$ *tangent to* $E$ *at* $x = 0$ *with the same dimension as* $E$;

*(ii)* $W(0)$ *is unique among all* $\mathcal{C}^{\sigma(E)+1}$ *invariant manifold satisfying (i);*

*(iii) If* $f_0$ *is jointly* $\mathcal{C}^r$ *in* $(x, \mu)$ *for some parameter* $\mu \in \mathbb{R}^p$, *then so is* $W(0)$, *this also holds for* $r = \infty$ *and* $r = a$.

*Proof.* See [Haller and Ponsioen, 2016] □

With this theorem, we can see that high-enough-order Taylor-series expansion puts us on $W(0)$ if we look for it as an invariant graph over $E$.

# Bibliography

[Arnold, 1992] Arnold, V. (1992). *Ordinary Differential Equations*. Springer.

[Breunung and Haller, 2019] Breunung, T. and Haller, G. (2019). When does a periodic response exist in a periodic forced multi-degree-of-freedom mechanical system? *Nonlinear Dynamics*, 98:1761–1780.

[Burns et al., 1999] Burns, T. J., Davis, R. W., and Moore, E. F. (1999). A perturbation study of particle dynamics in a plane wake flow. *Journal of Fluid Mechanics*, 384:1 – 26.

[Chicone, 1999] Chicone, C. (1999). *Ordinary Differential Equations with Applications*. Springer.

[Fenichel, 1979] Fenichel, N. (1979). Geometric singular perturbation theory for ordinary differential equations. *Journal of Differential Equations*, 31:53–98.

[Guckenheimer and Holmes, 1990] Guckenheimer, J. and Holmes, P. (1990). *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. Springer, New York, 3. print., rev. and corr. edition.

[Haken, 1977] Haken, H. (1977). *Synergetics: an Introduction, Nonequilibrium Phase Transitions and Self-organization in Physics, Chemistry, and Biology*. Springer Berline, Heidelberg.

[Haller, 2001] Haller, G. (2001). Distinguished material surfaces and coherent structures in three-dimensional fluid flows. *Physica D: Nonlinear Phenomena*, 149(4):248–277.

[Haller and Ponsioen, 2016] Haller, G. and Ponsioen, S. (2016). Nonlinear normal modes and spectral submanifolds: existence, uniqueness and use in model reduction. *Nonlinear Dynamics*, 86:1493– 1534.

[Haller and Sapsis, 2008] Haller, G. and Sapsis, T. (2008). Where do inertial particles go in fluid flows? *Physica D: Nonlinear Phenomena*, 237(5):573–583.

[Kutz et al., 2015] Kutz, J. N., Fu, X., and Brunton, S. L. (2015). Multi-resolution dynamic mode decomposition. *SIAM Journal on Applied Dynamical Systems*.

[Lotka, 1925] Lotka, A. (1925). *Elements of Physical Biology*. Williams and Wilkins Company.

[Maxey and Riley, 1983] Maxey, M. R. and Riley, J. J. (1983). Equation of motion for a small rigid sphere in a nonuniform flow. *The Physics of Fluids*, 26(4):883–889.

[Milnor, 1965] Milnor, J. (1965). *Topology from the differentiable viewpoint*. University Press of Virginia.

[Palis, 1967] Palis, J. (1967). *On Morse-Smale Diffeomorphisms*. PhD thesis, UC Berkely.

[Palis, 1969] Palis, J. (1969). On morse-smale dynamical systems. *Topology*, 8:385–404.

[Rouche et al., 1977] Rouche, N., Habets, P., and Laloy, M. (1977). *Stability Theory by Liapunov's Direct Method*. Springer New York, NY.

[Saari and Urenko, 1984] Saari, D. G. and Urenko, J. B. (1984). Newton's method, circle maps, and chaotic motion. *The American Mathematical Monthly*, 91(1):3–17.

[Sapsis and Haller, 2010] Sapsis, T. and Haller, G. (2010). Clustering criterion for inertial particles in two-dimensional time-periodic and three-dimensional steady flows. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 20(1):017515.

[Schmid, 2010] Schmid, P. (2010). Dynamic mode decomposition of numerical and experimental data. *Journal of Fluid Mechanics*, 656:5 – 28.

[Strogatz, 2000] Strogatz, S. (2000). *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry and Engineering*. Studies in nonlinearity. Westview.

[Sun et al., 2016] Sun, P., Colagrossi, A., Marrone, S., and Zhang, A. (2016). Detection of lagrangian coherent structures in the sph framework. *Computer Methods in Applied Mechanics and Engineering*, 305:849–868.

[Verhulst, 1989] Verhulst, F. (1989). *Nonlinear Differential Equations and Dynamical Systems*. Springer.

[Volterra, 1926] Volterra, V. (1926). Variazioni e fluttuazioni del numero d'individui in specie animali conviventi. *Memor. Accad. Lincei.*, 6:31–113.

[Wiggins, 1994] Wiggins, S. (1994). *Normally Hyperbolic Invariant Manifolds in Dynamical Systems*. Springer New York, NY.

[Williams et al., 2015] Williams, M., Kevrekidis, I., and Rowley, C. (2015). A data-driven approximation of the koopman operator: Extending dynamic mode decomposition. *Journal of Nonlinear Science*, pages 1307–1346.